DOCUMENT RESUME

ED 413 764 FL 024 836

AUTHOR Taura, Hideyuki
TITLE A Test on a Bilingual Dual Coding Hypothesis in Japanese-English Bilinguals.
PUB DATE 1996-12-00
NOTE 79p.; Master's Thesis, Macquarie University, Sydney, Australia.
PUB TYPE Dissertations/Theses - Masters Theses (042) -- Reports - Research (143)
EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Age Differences; *Bilingualism; Contrastive Linguistics; English; Foreign Countries; High School Students; High Schools; *Japanese; Language Research; *Language Role; Linguistic Theory; Recall (Psychology); Second Language Learning; *Second Languages; Sex Differences; Time Factors (Learning); Translation
IDENTIFIERS *Dual Coding Theory; Japanese People

ABSTRACT

A study investigated the effects of second language (L2) acquisition age, length of L2 exposure, and gender on bilingual coding, and examined whether the bilingual dual coding effect in incidental recalls would be the same as in Indo-European languages. The bilingual dual coding hypothesis proposes that the individual's image system and the two verbal systems corresponding to the two languages in question are functionally independent but partly interconnected. Subjects were 47 female and 17 male balanced Japanese-English bilingual high school students at an Osaka (Japan) school for children returning from periods abroad. They were presented with pictures to be labeled in English, Japanese words to be translated into English, and English words to be copied. Later, without warning, they were tested to recall the generated English words. Results support the bilingual dual coding hypothesis, with no gender effects observed. Onset age of L2 acquisition was found to be significant. It is concluded that the theory is generalizable across bilinguals in both Indo-European alphabetic languages and in this alphabetic/non-alphabetic language combination, and also in child and well as adult bilinguals. Contains 97 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

A Test on a Bilingual Dual Coding Hypothesis in Japanese-English Bilinguals

ED 413 764

by

Hideyuki Taura

Submitted as partial fulfilment of the requirements
for the degree of Master of Applied Linguistics,
Macquarie University, Sydney, Australia.

December, 1996.

I declare that this work has not been submitted
for a higher degree at any other university or institution.

Signed .....

Data .....

FL024836

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Hideyuki Taura

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

ABSTRACT

This study is to (1) investigate the effects (if any) of L2 acquisition age, the length of exposure to L2, and gender on the bilingual dual coding hypothesis proposed by Paivio and Desrochers (1980) and (2) verify whether the bilingual dual coding effect in incidental recalls would be the same in Japanese-English as in Indo-European languages.

Balanced Japanese-English bilingual subjects are presented with (1) pictures to be labelled in English, (2) Japanese words to be translated into English, and (3) English words to be copied as they are. Later without warning they are tested to recall the generated English words.

The results showed a 3.7 : 3.2 : 1.0 ratio for pictorial : translation : copy encoding conditions, which is supportive of the bilingual dual coding hypothesis. Both small pictorial-translation difference and the high ratio for translation were interpreted as caused by the Japanese language specific effects - logographic features. No LOR or gender effects were observed. The onset age of L2 acquisition proved to be a significant factor, which added an extension to Arnedt and Gentile's (1986) 'manner' - proficiencies of L1 and L2 prior to formal schooling as well as the language sequence in schooling (L1 to L2 or L2 to L1) should be considered in the bilingual dual coding framework. Thus, this experiment seems to reveal that the bilingual dual coding hypothesis is generalizable across (1) bilinguals not only in Indo-European alphabetic languages, but also in alphabetic/non-alphabetic languages, and (2) bilinguals in their childhood as well as adulthood.

CONTENTS

	Page
LIST OF FIGURES AND TABLES	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
1. INTRODUCTION	
1.1. Bilingual memory	1
1.2. Hybrid models	2
1.3. The bilingual dual coding model	6
1.4. Tests on the bilingual dual coding model	9
2. METHODOLOGY	
2.1. Subjects	21
2.2. Measurement of balanced bilinguality	21
2.2.1. Cloze- / C- tests	23
2.2.2. Self-rating of language proficiencies	26
2.2.3. Teacher judgements	27
2.2.4. Selected balanced bilingual subjects	28
2.3. Materials	29
2.4. Procedures	29
3. RESULTS	33
4. DISCUSSION	37
4.1. Gender	37
4.2. LOR	39
4.3. Onset age of L2 acquisition	47
4.4. Language specificity	56
5. CONCLUSION	62
REFERENCES	
APPENDICES	

List of Figures and Tables

Figures

1.1. Models by Paradis and Snodgrass

1.2. Bilingual dual coding model by Paivio

3.1. Incidental free recall scôres

4.2. LOR and mean recalled words

4.3. Mean numbers of recalled words

Tables

2. Balanced bilingual subjects selection

3.1. Mean words recalled for treatment conditions by subgroups

3.2. Interaction effects on 3 variables and encoding treatment

3.3. Paired t-Tests on 2 onset age groups and encoding treatment

4.2. LOR

4.3.1. Onset age and number recalled words in each condition

4.3.2. Comparison of ratio and difference of ratio in three experiments

4.3.3. Comparison of mean age in three experiments

4.4. Comparison of ratios, differences and languages

1. INTRODUCTION

1.1. BILINGUAL MEMORY

A controversial issue in research on bilingual memory concerns whether two languages are stored interdependently or independently. Proponents of the interdependence view argue that bilingual code-switching can be satisfactorily explained in the unitary framework in that both languages are concurrently active and interactive on the lexical and syntactic levels. Swain (1972, cited in Paradis and Lebrun 1984, p32) rejects the independent model, stating that two separate lexicons for both languages are not efficient and economical in terms of memory storage. A number of other researchers support the common-storage model with whole ranges of empirical data (Kolers, 1966a; Dalrymeple-Alford and Aamiry, 1969; Lambert, Ignatow, and Krauthamer, 1968; Nott and Lambert,1968; Macleod, 1976; Caramazza and Brones, 1980).

Almost an equal number of studies, however, have also provided evidence for the independence view (Taylor, 1971; Kolers and Gonzalez, 1980; Kirsner, Brown, Abrol, Chadha, and Sharma, 1980; Scarborough, Gerard, and Cortese, 1984; Cristoffanini, Kirsner, and Milech, 1986). The separate-storage proponents have gained evidence from the experiments which elicit significant response in one language independent of the other or additive effects of the two languages in memory tasks.

1.2. HYBRID MODELS

Drawing on these mixed data which are indeterminate regarding the distinction, recent research suggests that both alternatives are correct (e.g. Snodgrass, 1995, 1993, 1985, 1984; Matsumi, 1994; Woutersen, De-Bot, and Weltens, 1995; Williams, 1994; De Groot, 1993; Kirsner, Lalor, and Hird, 1993; Paivio, 1991, 1986, 1983; Paivio, Clark, and Lambert, 1988; Durgunoglu and Roediger, 1987; Potter, So, von Eckhardt, and Feldman, 1984; Paradis, 1978,1981; Paivio and Lambert, 1981; Paivio and Desrochers, 1980). The empirical discrepancies are interpreted as there being no need to differentiate the unitary from the separate model, as various types of processing could take place according to strategies used by bilinguals such as attitudes, skills, goals, and purposes (Hoffmann, 1991 P78; Durgunoglu and Roediger, 1987 p379). In other words, bilinguals may use shared representations across their two language, but the conditions under which they can do so are constrained by various factors such as the nature of the material used (concrete vs. abstract), the form of elicitation method, the level of proficiency in the second language, the availability of lexical-level connections between the two languages, and the strategies used by bilinguals (Kroll, 1993, p53; Hummel, 1993, p270). Contradictory results, therefore, emerge as a result of the variation in each experiment in terms of which aspects of bilingual processing researchers choose to investigate, as well as the control and intervening variables.
Researchers have put forward several models in support of the hybrid argument.

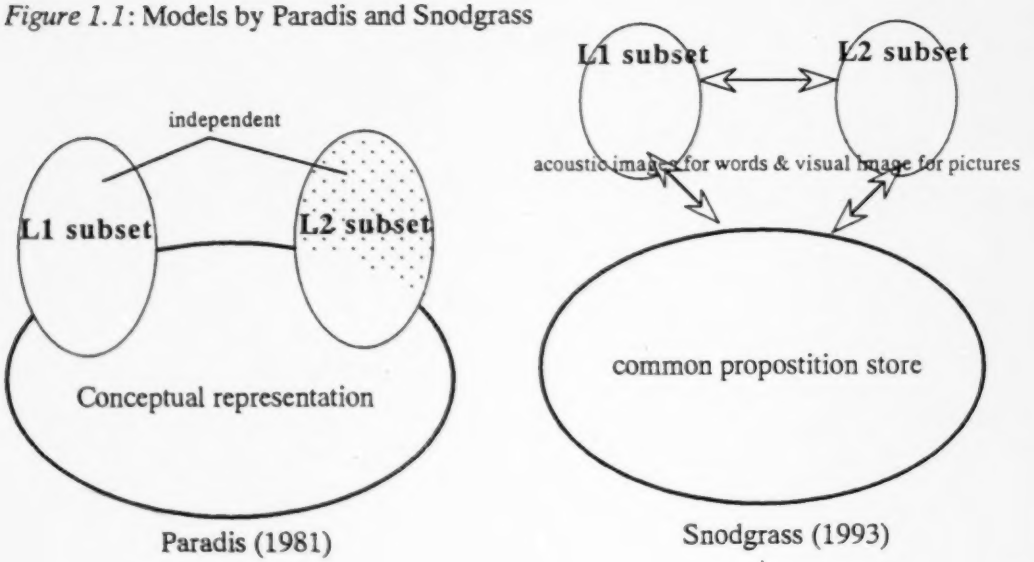
Kolers, for example, based on the data collected from a series of

word-association experiments (e.g. 1963, 1965, 1966, 1968), concluded that the degree of semantic overlap of two languages in a bilingual is not the same for concrete and abstract words. Clark (1978, cited in Hamers and Blanc, 1989, p95) and Hammound (1982, cited in Hamers and Blanc, 1989, p95) have reported similar results, which showed that bilinguals seem to have a more shared organization of concrete words and a more separate one for abstract words. These researchers argue that the concrete-abstract distinction could satisfyingly account for the discrepant results of the studies varying in empirical paradigms. According to this theory, the studies which involve tasks with concrete words are apt to produce results interpretable as supporting a common storage, whereas those involving tasks with abstract words are likely to result in evidence for a separate model.

Paradis (1978,1981) proposes that although two languages are stored in a single system, there may be elements of each language which form subsystems within the larger system (Figure 1.1.). Thus, bilinguals have two subsets of neural connections, one for each language, which can be activated or impeded independently. At the same time, they share a general conceptual store from which they are able to select elements of either language at any time. Very similar to this is the hybrid model proposed by Snodgrass (1995, 1993) in Figure 1.1. According to this framework, separate stores exist at a shallow level of processing containing both acoustic and visual images, however at the higher level there is a common proposition storage. These two models seem to indicate that the greater number of conceptual features shared by a word and its translation equivalent in the other language, the more

they will tend to evoke the same response in recall or association tasks. Therefore bilingual association-tasks which use words conceptually shared by both languages are expected to produce data supportive of a unitary model, while studies involving words stored in language-specific subsets are likely to support an independent model of bilingual lexicons.

Figure 1.1: Models by Paradis and Snodgrass



While Snodgrass's hybrid model has two separate subsets for each language which store acoustic and visual images, a third model proposed by Paivio and Desrochers (1980), Paivio and Lambert (1981), and Paivio, 1986, 1991) postulates one common image system which is referentially connected to concrete word logogens in individual verbal systems as shown in Figure 1.2. This third approach to the study of

bilingual memory is dealt with in more detail and tested in the present research.

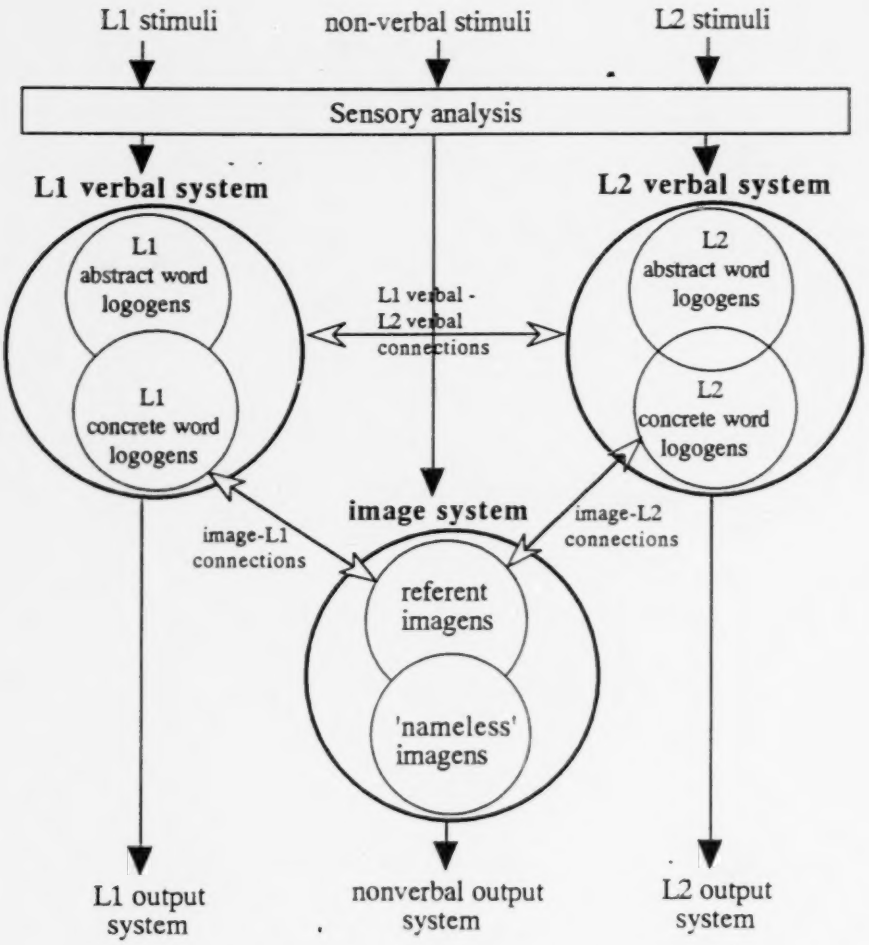


Figure 1.2: Schematization of the bilingual version of dual coding system (Source: Paivio, A., 1991. *Images in mind* . p329)

* 'Logogen' is a term used by Morton (1979) to describe a unit of verbal representation which is a process by which words become accessible in the presence of stimuli. Paivio (1981) uses the term as a cognitive process which generates a word, while 'imagen' is used as generating an

evoked image.

1.3. THE BILINGUAL DUAL CODING MODEL

The model proposed by Paivio et al. employs a bilingual dual coding system. This differs from the previous models in that the imagery system, which represents concepts - knowledge of the world - and the two verbal systems corresponding to each language of the bilingual are all independent but partly interconnected (knowledge of the world is available to the both verbal systems).

The two verbal systems and the image system are functionally independent in the sense that encoding, storing, and retrieving non-verbal objects/events is possible without any intervention from either of the verbal systems and, conversely, that bilinguals can engage in verbal activities without being influenced by the image system. As for the two verbal systems, one verbal system can function in terms of comprehension, memory, and production independent of the other verbal system. Code-switching occurs when the input language changes or contextual cues for a code switch are delivered by interlocutors. Once code-switching is triggered, language processing proceeds within a system based on the rules specific to the language in question without any intervention from the other verbal system (Paivio, 1991, p330).

Interconnectedness, on the other hand, implies that the image system can influence verbal behaviour in either language system at the

referential level and vice versa. Regarding the L1-L2 verbal system relations which are connected through the translation path, the translation equivalents in L1 and L2 do not necessarily activate identical referential meanings (e.g., the translation equivalents English 'bread' and French 'pain' could arouse images of different kinds of bread as Arnedt & Gentile (1986, p291) mention). This is partially due to the referential connections to the image system from the two verbal systems. These systems share some connections or are partly independent according to how the two languages are acquired. Bilinguals, for instance, who have acquired two languages in bilingual surroundings which require constant reference to the same things in two languages, are likely to have stronger connections between translation equivalents. In contrast, individuals who have learnt two languages in totally independent settings where they have little or no need to translate, would then find it difficult to translate. Another explanation for activation of non-identical referential meanings by the translation equivalents in L1 and L2 is that images are known to be determined by the context as well as the abstractness-concreteness of the underlying concept (Bugelski, 1970; Paivio, 1965, 1971).

This argument that the bilingual's two verbal systems are functionally independent but partly interconnected leads Paivio and Desrochers (1980) to discuss further implications. Firstly, while translation primarily takes place directly through the L1-L2 connections, indirect translation via the image system is also possible under some circumstances: an L1 logogen activates imagens which in turn induce the activation of L2 logogens allowing the translation

equivalents to be accessed. This could be tested by an experiment designed to examine additive retention effects in comparing a modality of simply presenting a word to be translated with that of presenting a word together with its picture for translation. According to this model, it is predicted that the picture-encoding modality would induce a higher retention rate than the other modality due to the increased number of L1-L2 connection ties via imagens.

The second implication concerns the stronger ties between L1-L2 translation equivalents than the associative connections within each language. The translation equivalents between L1 and L2 are presumed to approximate one-to-one relations (e.g. boy-garcon in English-French bilinguals) are more certain and constrained because of frequent reference in translation. Associative connections within each system, on the other hand, involve one-to-many (e.g., boy-girl, man, and father etc), and they are rich and diverse because they result from varied verbal contexts and experiences (Paivio, 1986; Harbluk, Paivio, and Clark, 1987; O'neill and Roy, 1993). This could be verified by greater free recall in translation than mere copying words which occurs only in a unilingual framework. Individuals vary depending on the strength and number of ties in between verbal systems or within each system.

This second implication can successfully explain the difference between coordinate and compound bilinguals, which was proposed by Weinreich (1968). The two types can be viewed as the poles on a

continuum. In 'compound' bilinguals, some connections from the two verbal systems are assumed to converge on a common set of imaginal representations in the image system whereas translation equivalents in the two verbal systems activate different sets in the imagens in 'coordinate' bilinguals.

Thus the bilingual dual coding hypothesis seems to provide a valid explanation for the discrepant results in studies on bilingual brains, including the issues of coordinate-compound distinction and how code-switching functions. In testing this model, a number of researchers have examined the additive retention effects in bilinguals from imagery, bilingual, and unilingual codings. They base their methodological reasonings on the assumption that picture cues, which would activate three systems (L1, L2 verbal systems and an imagery system), will produce a higher retention (or recall) rate than translation cues, which would activate only two verbal systems, whereas translation cues will produce a higher retention rate than copying cues, which would activate only one verbal system.

1.4. TESTS ON THE BILINGUAL DUAL CODING MODEL

The earliest research to test the model was conducted by Paivio and Lambert (1981). In their first experiment, 12 balanced French-English bilingual university students were visually presented with pictures, French words, or English words (17 of each) with the instructions to label the pictures in English, translate the French words into English, or simply copy the English words. Later without prior

warning the subjects were asked to recall the generated English words in any order. In the second experiment, the presentation modality was reversed to counterbalance: English stimulus words were presented to be either sketched (the subjects were asked to form a mental image of the word's referent and then sketch the item referred to), translated into French, or copied. Retention in both experiments was best for the imagery-based condition, next for the translation-based condition, and last for the copying condition in a ratio of approximately 3:2:1. Paivio and Lambert concluded these additive effects to be completely consistent with the bilingual dual coding hypothesis, which predicts that recall is highest for pictorial coding (which activates English and French verbal systems as well as an imagery system), intermediate for bilingual-coding (which activates 2 verbal systems), and lowest for the unilingual copy condition (which activates only 1 verbal system), and that the overall recall ratio would be 3:2:1 according to the number of systems activated by cues.

This pattern of results has been replicated in a number of subsequent studies (Paivio, 1983, 1986, 1991; Watkins and Peynircioglue, 1983; Paivio, Clark, and Lambert, 1988; Vaid, 1988; Matsumi, 1994). As well as testifying the additive effects in the three different codings, the subsequent studies have also managed to extend Paivio and Lambert's (1981) findings to generalize across different parameters such as varied stimulus exposure durations, words and pictures presented, and levels of second language proficiency.

Exploring the issue of L2 acquisition history, or the age of L2 acquisition, Paivio and Lambert (1981) compared the early (infant) bilinguals and late (childhood or early adolescence) bilinguals and found no early-late bilingualism effects in Experiment 1. However, when the subjects were presented with a list containing the English version of the 51 items which they had been shown, and asked to indicate in what mode each words was presented in Experiment 1, late bilinguals recalled presented mode better than early bilinguals. Furthermore in Experiment 2, the superiority of the image over translation conditions was not significant for late bilinguals. The researchers did not interpret the data any further than merely referring to as 'having potential theoretical significance but only suggestive since they did not emerge in the subject analysis' (p537) and 'not being discussed further here since the interaction did not occur in the subjects analysis and it does not seriously qualify the overall pattern of results' (p538).

Arnedt and Gentile (1986), who took this point, criticized Paivio and Lambert's failure to offer any specific prediction as to L2 acquisition history in the bilingual dual coding framework. They interpreted Paivio and Lambert's discrepant data in the following manner. When late bilinguals start to acquire L2, they predominantly use translation connections between L1 and L2 verbal systems at an initial stage. They gradually develop numerous and more strengthened connections between the L2 verbal system and the image system, until they become balanced bilinguals. Therefore for late bilinguals the translation connections between L1-L2 systems are stronger than those

between L2-image systems. Early bilinguals, on the other hand, develop their L2-image connections as strongly as their L1-image systems, and they have no need to strengthen the L1-L2 translation ties as long as they acquire the two languages in independent settings, which is often the case with early bilinguals. Thus Arnedt and Gentile interpreted that late bilinguals' L1-L2 connections are stronger than early bilinguals' and their L2-image ties are not as strong as early bilinguals' ties.

This leads to the logical explanation of the discrepancy in Paivio and Lambert's results mentioned above. Arnedt and Gentile argue that early bilinguals are expected to have 'an image-rich second language environment' (p293) while late bilinguals are less likely to be in such an environment, and therefore their imagery-translation differences should be reduced and L1-L2 translation ties should be enhanced. This theoretical framework seems to satisfactorily explain Paivio and Lambert's discrepancy that the superiority of the imagery over the translation conditions was not significant for late bilinguals.

Thus Arnedt and Gentile point out the manner of L2 acquisition, not the age of L2 acquisition as the relevant variable which caused the discrepancies in imagery-translation, rather than imagery-copying or translation-copying. In order to operationalize this theory and test its validity, in their experiment (1986), Arnedt and Gentile used a pool of bilingual subjects who varied in manner of L2 acquisition. They predicted that bilinguals who study in immersion settings (similar to naturalistic environments) should develop larger imagery-translation

differences in memory, while those who are in traditional translation programs should show smaller imagery-translation differences. Thus they expected a larger imagery-translation differences in the immersion group than the traditional group. The results supported their predictions showing 3.9 : 2.4 : 1 for the immersion group in the pictorial, bilingual, and unilingual coding and 3.1 : 2.4 : 1 for the traditional group. The manner of L2 acquisition, therefore, was assumed to affect the strength of connections among L1-L2-image systems.

This experiment by Arnedt and Gentile, however, does not provide unequivocal evidence to reject the idea that the age of L2 acquisition affects the connection ties. They failed to present the L2 acquisition ages of their subjects to statistically show that acquisition age plays no role in the bilingual dual coding framework. No subsequent research seems to raise the L2 acquisition issue with sufficient age information provided. Therefore, one of the present study's purposes is to statistically analyse the L2 acquisition age effects (if any) on the dual coding theory by using bilinguals who have leant L2 only in natural settings (equivalent to Arnedt & Gentile's immersion settings) to eliminate a latent intervening variable.

In distinguishing the subjects into early and late bilinguals, Paivio and Lambert (1981) did not explicitly specify the dividing age (nor did Matsumi, 1994). According to Paivio and Descrocher's study (1980) where Lambert's notion of early/late bilingual distinction is cited (p332), an early bilingual is defined as 'those brought up in a thorough

bilingual environment from infancy' and late as 'some time after infancy, usually after ten years of age'. Some studies on the maturational constraints of L2 acquisition seem to converge on two dividing ages, depending on what aspects of language (i.e. phonology or syntax etc.) are examined; Oyama (1976, p261) age 12 for phonology, Newport (1990) 4, 6, and 12 for American Sign Language, Johnson & Newport (1991) 4-7 for universal grammar, Scovel (1988, p153) 12 for neuroplasticity, and Palij (1990) 6 & 12. In line with these results, particularly with Palij (1990), who looked into bilingual L2 acquisition by grouping the subjects according to their onset age of 0-6, 6-12, and over 12 as well as Hamers and Blanc (1989, p9), who define child bilinguality as acquired before age 11 and adolescent bilinguality between 11 & 17, the present study initially set 2 onset age dividing lines, that is, age 6 and 12. However, due to the age range of the subject group (onset age range of 0-16, but only 5 subjects out of 64 are over onset age 12), this study draws a single dividing line of onset age 6 to explore the effects of onset age on the groups 0-6 and over 7 in the bilingual dual coding theory framework.

Effects of length of residence (LOR, henceforth) in L2 language speaking communities or countries are also examined. LOR is taken into account based on the results by a number of researchers who have found significant effects of LOR on L2 acquisition in bilinguals (e.g. Cummins, 1984, 1991; Krashen, 1982; Romaine, 1995; Yumoto, 1995; Ono, 1994; Walberg, Hase, and Rasher, 1978 cited in Krashen, 1982). Cummins (1991), for instance, who examined Asian students (Chinese

and Japanese) in North America, maintains that LOR is a major factor in attaining English skills. Accordingly, this study examines LOR effects by dividing the subjects into 3 groups depending on their LOR: the first group of LOR 0 - 4 years, the second 5 - 10 years, and the third over 11 years. This distinction was made based on the results of the following research. Firstly, investigating Japanese children who are living in English-speaking countries, Ono (1994, p100) concludes that it seems to take about 3 to 5 years (depending on their onset age of L2) for them to academically catch up with the monolingual counterparts in their host countries. Secondly, in their study with Japanese immigrant children in Toronto, Harley, Allen, Cummins, and Swain (1990) found that at least four years were required for students from highly educated backgrounds to attain grade norms on English academic tasks. Thirdly, Mgiste (1992), who compared the developmental changes in picture naming and number naming of 151 German immigrants to Sweden (aged 6-19), argues that after about four years of LOR the subjects reached language balance on the task. Finally, citing Walberg, Hase, and Rasher's study (1978), which examined Japanese-speaking children in the United States and found a significant relationship between LOR and English proficiency, Krashen (1982, p39) mentions a maximum LOR, beyond which there is no relationship between LOR and L2 acquisition. Walberg et al were cited as 'acquisition proceeds at a fast rate initially, but the amounts of gain diminish with time..... units are gained in the first two months, the following two years, and the next eight years'. Thus it seems reasonable to categorize the subjects into 3 groups according to LOR - (1) 0 - 4, (2) 5 - 10, and (3) over 11 years.

With respect to gender, Paivio and Lambert's Experiment 1 (1981), where an equal number of male and female subjects were tested, showed no effects for gender either alone or in interaction with other conditions. In the postexperimental test on recalling the presentation mode, however, males surpassed females in accuracy. Similar to the early-late bilingual issue as mentioned above, the researchers declined to interpret the data with more than a suggestive comment that 'these results have potential theoretical significance, however they can only be viewed as suggestive since they did not emerge in the subject analysis' (536-537). The researchers seem to regard the additive effects observed in Experiment 1 as significant, and the postexperimental test as unimportant since the latter only deals with accuracy of recalling the presentation mode, not with the retention itself.

Despite their rigorous re-examination of Paivio and Lambert's study to discover significant factors in the way of L2 acquisition, Arnedt and Gentile (1986) too, did not seem to find any significance in the gender issue. Ignoring the gender issue, male subjects accounted for only 7 % in the Arnedt and Gentile experiment. In order to clarify this gender issue, the present research examines interactions (if any) between gender and the additive effects in the pictorial, bilingual, and unilingual modalities.

In previous studies testing the bilingual dual coding hypothesis, subjects were bilingual or multilingual mostly in Indo-European

languages such as (1) French and English (Paivio and Lambert, 1981; Arnedt and Gentile, 1986; Paivio, Clark and Lambert, 1988; O'Neill, Roy and Tremblay, 1993), (2) Dutch, English and French (Groot & Hoeks, 1995), (3) Afrikaans and English (Tyson, Doctor, and Mentis, 1988), and (4) Spanish and English (Kolers and Gonzalez, 1980; Scarborough, Gerard, and Cortese, 1984). Matsumi (1994), however, investigated the bilingual dual coding hypothesis, using Japanese-German early balanced bilinguals, Japanese-English late balanced bilinguals, and Japanese learners of English as a second language. In this study, the subjects were required to encode a mixed list of L1 and L2 words by translating L1 words and copying L2 words with either imagery or non-imagery instruction. In the imagery condition, the subjects were instructed to generate an image in response to a stimulus both by copying L2 words and in translating L1 words into L2. Without prior warning, the subjects were then asked to recall L2 words. The results showed an approximately 3:2:1 ration for item recall in the imagery condition, the translation with non-imagery condition, and the copying with non-imagery conditions. The additive effects, however, were only partially observed with the Japanese learners of English as a second language. He concluded that the bilingual dual coding theory is possible for balanced bilinguals, whereas the theory is not applicable to second language learners who have not developed sufficient interconnections between the second language system and the imagery system.

Matsumi's experiments, however, seem problematic in two respects when comparing his results with those from other bilingual

dual coding studies. Firstly, the Japanese subjects in his three experiments were not consistent in their use of L2, that is, German in the first experiment and English in the second and third experiments. Mentioning his inaccessibility to late-Japanese/German bilinguals and Japanese learners of German as a second language, Matsumi justified his experiments by (1) claiming that both German and English belong to the same German family in Indo-European languages, and (2) keeping constant the average ratings of words in each language for their ability to elicit an image by native speakers. Eliminating possible intervening variables would be undoubtedly desirable when comparing the results. Therefore the present study uses the subjects whose L2 is kept the same - English.

Secondly, although Paivio & Csapo (1973) and Snodgrass & McClure (1975) verified that generating an image in response to a stimulus word produces the same effect as presenting a picture, their studies were conducted in a unilingual framework which only engages the same language coding system. In order to compare the direct relationship between the results, it would be desirable that elicitation methods be kept the same across the studies. Therefore the present study employs the same picture-labelling elicitation method as the previous bilingual dual coding experiments when it investigates whether the bilingual dual coding hypothesis could be explained in languages where diverse writing systems such as English and Japanese are used.

The Japanese language is different to the major Indo-European languages in eliciting image from words, therefore a brief explanation of the language is provided before furthering the discussion. The Japanese language uses three different forms of characters in writing: *kanji* (Chinese characters), which is logographic, and two kinds of syllabaries, *hiragana* and *katakana,* both of which are phonographic (Hirose, 1992. p.908). *Kanji* is typically used for nouns, root parts of verbs, adjectives, and adverbs, while *hiragana* provides inflections and other grammatical parts of sentences, and *katakana* gives a transcription of foreign loan words (Cabeza, 1995. p.156).

Regarding the logographic feature of *kanji,* Chen and Juola (1982) argue that languages using a logographic writing system seem to activate encoding strategies which emphasize visual codes, whereas a visual strategy is relatively less important than is a phonological one in languages which use alphabetic writing system such as English. A similar report was made by Tzeng and Wang (1983) in their claim that processing a logographic writing system involves more visual memory than does processing alphabetic scripts. They found that Chinese readers recalled better when materials were presented visually rather than auditorily but no such difference was found for English readers. In the same vein, Paivio (1986) suggests that a high-imagery language is understood by visualizing its semantic content while a low-imagery language is understood by its intraverbal patterning.

Synthesizing above mentioned argument, Ho and Chen (1993, p509) hypothesized that 'as both the processing of imagery and that of a logographic system emphasize visual processing strategies, imagery

could facilitate memory coding more for logographic language users than for alphabetical language users'. Accordingly the present study, using Japanese/English balanced bilinguals, predicts that if the bilingual dual coding hypothesis is to have generalizability to the population at large (i.e., bilinguals not only in same or similar language families but also languages from different language families), then the 3:2:1 magnitude should be obtained. It further predicts that the recall ratio for Japanese words to be translated into English would be closer to the ratio for pictures to be labelled in English because of the logographic features of *kanji* as long as the majority of Japanese words presented in this study are represented by *kanji*.

In sum, this study is designed to verify the bilingual dual coding hypothesis in terms of effects caused by the following four factors:

(1) onset age of L2 acquisition and (2) LOR

The present study predicts that the earlier and longer bilinguals are exposed to L2 in natural settings, the more strengthened the imagery-L2 connection ties are. This working hypothesis will be verified by the results that bilinguals who have acquired L2 earlier in their lives and who have stayed in L2 speaking communities for a longer period of time, display a larger imagery-translation difference in memory than those who have acquired L2 later and whose LOR is shorter.

(3) gender

The gender issue is explored to clarify whether or not it plays any role in the bilingual dual coding theory.

(4) language-specificity

This study examines the applicability of the bilingual dual coding theory to bilinguals of Roman alphabetic and non-Roman alphabetic

languages. It is predicted that the recall ratio for Japanese words for translation would be closer to that for pictures labelled in English, due to the logographic feature of Japanese *kanji*.

2. METHODOLOGY

2.1. SUBJECTS

Subjects are 64 (17 male, 47 female) balanced Japanese/English high school 'returnee' students (aged 12 - 18 , mean age of 15.8) at Osaka Intercultural Academy (OIA) in Osaka, Japan (see 2.2.4. for balanced bilinguals selection procedures). OIA is a private high school which was established in 1991 to mainly accept the ever-increasing Japanese 'returnee' students from overseas (approximately 50,000 school-aged Japanese children are reported to be living outside Japan as of May 1, 1995; the Japan Times June 24, 1996). OIA's faculty and facilities are shared with Osaka International School, which creates an environment for the students to maintain and improve their two languages simultaneously. Every subject in the experiment has been exposed to English speaking societies (40 students in U.S.A., 7 in England, 4 in Australia, 4 in Canada, 1 in New Zealand, 7 in international schools outside Japan, and 1 in a bilingual family in Japan) and all learned English as an infant or a child. Since at least one of their parents is Japanese with 3 subjects and both parents are Japanese for 61 subjects, Japanese has been used all their lives even when they were living outside Japan.

2.2. MEASUREMENT OF 'BALANCED' BILINGUALITY

Special attention was drawn to the measurements of balanced bilinguality in this research, for the bilingual dual coding theory is

constructed in a framework of fluent or balanced bilinguals (Paivio and Desrochers, 1980; Vaid, 1988; Matsumi, 1994). Therefore the definition of bilinguality and the measurements of balanced bilinguality are examined before furthering the discussion.

According to Hamers and Blanc, bilinguality is "a psychological state of the individual who has access to more than one linguistic code; the degree of access will vary along a number of dimensions" (1989, p6). Six dimensions relevant to bilinguality are listed by them: relative competence, cognitive organization, age of acquisition, exogeneity, social cultural status and cultural identity. Thus bilinguality is a multidimensional phenomenon and it should be measured not only by the simple sum of two monolingual competences, but also by evaluation of unique bilingual characteristics such as code-switching. However, due to the present lack of sufficient methods to capture the specificity of bilingual behaviour, it is the general practice to use a variety of measures to grasp a state of bilinguality (Hamers et al, 1989).

Therefore, in assessing 'balanced' bilinguality in the studies of the bilingual dual coding theory, researchers have used various methods. Lambert and Rawlings (1969) used a self-rating scale, Paivio and Lambert (1981) used a language background questionnaire and a bilingual version of the Stroop colour-word task, Arnedt and Gentile (1986) used a language biographical data form, Vaid (1988) used self-ratings of proficiency, Paivio, Clark and Lambert (1988) used self-ratings, interview with native speakers to judge verbal fluency, and a bilingual Stroop test, and Matsumi (1994) used a bilingual Stroop test

and self-evaluation in each language. With reference to these methods, the present research applied the following three measures to assess the global balanced bilinguality of the subjects: (1) Cloze-/C-tests, (2) self-evaluation of language proficiency, and (3) teacher judgements. A bilingual version of the Stroop colour word test was not included for the technical reason that the device to measure time was unavailable.

2.2.1. CLOZE-/C-TESTS

Cloze-tests have been used to assess the relative proficiency of bilinguals in two languages on the ground that they are good tools to measure the unitary nature or global view of language proficiency (Oller, 1979; Cohen, 1980, Hamers and Blanc, 1989). One of the reasons is that the tests require examinees to simultaneously utilize various cues such as grammatical, syntactical, lexical, semantic, collocational, pragmatical, logical, and situational cues (Klein-Braley, 1985). The reliability of the tests is supported by their high correlation rate with TOEFL .83 (Oller, 1979) and with the US Foreign Service Institute (FSI) Oral Interview (Cohen, 1980), which is regarded as one of the most reliable measurements of global communicative competence (Shohamy, 1983). Furthermore, Chihara and Oller (1978, cited in Krashen, 1982, p35) in studying Japanese learners of English (the same target population as the present research), used a cloze test along with other measures and supported its reliability.

Emphasizing the differences between cloze-tests and C-tests which seem similar at a surface level, some researchers (e.g. Piper, 1983;

Klein-Braley, 1985; Hughes, 1989; Hood, 1990; Singleton and Little, 1991) argue that C-tests have an advantage over cloze-tests in examining the bilingual L2 lexicon. C-tests, which delete the second half of every second word after the lead-in, instead of deleting whole words, are adopted as one of the placement test items by Piper (1983) with the result that they correlate with teacher judgements at 0.76 and are more reliable with the advanced language groups. The present research, having subjects with high proficiency in English, used an English C-test taken from Piper (1983, Appendix 1), which was originally used for Dutch students with the same age range of 13-18 as the subjects in this study.

With no Japanese C-tests yet designed due to technical difficulties where the continuous structure of the Japanese language is broken only by one-syllable grammar particles that cannot be deleted in C-tests, the researcher constructed cloze-tests for Japanese proficiencies (Appendix 2). Strictly following the guidelines on cloze-test construction by Klein-Braley (1985), Oller (1979) and Cohen (1980), the test material was based on a passage from a Grade 6 Japanese textbook screened by the Japanese Ministry of Education with every 5th word deleted regardless of content or function words.

In scoring the C-/cloze-tests, the appropriate word method was used over exact word method. Although the two methods correlate very highly (e.g., r=0.97; Cohen, 1980), the acceptable method is generally

adopted (Heaton, 1988; Madsen, 1983; Cohen, 1980; Hood, 1990). Accordingly this research was lenient with misspellings, apostrophes, singular-plural differences, and equivalent words against the normatives obtained from monolinguals (English normative data were taken from twenty 9th graders at Sandgate State High School in Queensland, Australia while Japanese data were drawn from twenty 9th graders at Osaka Intercultural Academy in Japan). With reference to assessing English proficiency, the subjects are expected to score more than 95% to be proficient since native speakers score nearly 100% in C-tests (Hughes, 1989, p71), which was supported by the data from Sandgate State High School samples showing that 19 students were marked at 100% and 1 student 96%.

As for cloze-tests, on the other hand, Heaton (1988, p31) argues that scoring over 53% indicates the material is easy enough for students to read independently while scores between 44 and 53% indicate the material is suitable for use at an instructional level, and scoring below 43 put students at a frustrational level. Referring to the scores of cloze-tests as an indicator of reading comprehension, Cohen (1980, p101) illustrates that 43% on a cloze-test is comparable to a score of 75% on a standard multiple-choice test of reading comprehension. Using both Heaton and Cohen's arguments would lead to the conclusion that 53% on the present Japanese cloze-test would prove sufficient proficiency, taking it into account that the cloze-test material was taken from Grade 6 textbooks and that the subjects are all 7th graders or older. However, due to the mean scores of 69% in the normative samples from OIA as well as the 95% set for the English C-test, at least 60% is required in the

present Japanese cloze-tests to claim sufficient proficiency.

The apparent difference between 95% for English and 60% for Japanese is due to the relative easiness of C-tests where the first half of each word is given, and the relative difficulty of cloze-tests as argued by Cohen (1980) above. In sum, the present study requires over 95% in the English C-test and over 60% in the Japanese cloze-test to claim proficient balanced bilinguality, on which bilingual dual coding theory is based.

2.2.2. SELF-RATING OF LANGUAGE PROFICIENCIES

The self-evaluation of proficiency in both languages is adopted in this research. This method has been found to correlate highly with more objective proficiency tests (Fishman & Cooper, 1969, cited in Vaid, 1988, p86; Durgunoglu and Roediger, 1987). In using Macnamara's 7-point scale, researchers (e.g. Lambert and Rawlings, 1969; Scarborough, Gerard, and Cortese, 1984; Schwanenflugel and Rey, 1986; Vaid, 1988; O'neill, Roy, and Tremblay, 1993) stipulated that the subjects have to rate themselves by at least 4 out of 7 (where 7 indicated proficient competence) in three of the four modalities to claim their balanced bilinguality. The version used in this study (Appendix 3) compares language capabilities both in English and Japanese in terms of speaking, reading, writing, and comprehension on a 5 point scale (5 being 'a lot better than monolinguals', 4 'slightly better', 3 'almost the same', 2 'slightly inferior to monolinguals', and 1 'a lot inferior') and specifies that the subjects have to score more than 3 points for at least 3 of the 4

skills to be claimed as 'balanced' bilinguals.

As for the reliability of self-assessment in terms of models against which the subjects measure their own language proficiencies, Romaine (1995, p16) cites Skutnabb-Kangas (1984),

> Swedish speakers' assessment of their knowledge of Finnish is more reliable in Finland than Finnish speakers' assessment of their knowledge of Swedish because the Swedish speakers have a greater chance of coming into contact with Finnish speakers and therefore have a model against which to measure their performance.

Attending OIA, the present subjects are in daily contact with both Japanese and English monolinguals. Therefore it is expected that self-assessment should be reliable in this respect.

2.2.3. TEACHER JUDGEMENTS

Language abilities of both English and Japanese in the 4 skills are assessed by the subjects' teachers. This assessment is carried out based on Hamers and Blanc's argument (1989, p21) 'evaluation of proficiency in both languages by native speakers of each language can be used as a reliable measure of balanced bilinguality'. Japanese language abilities are assessed by the subjects' current Japanese native instructors in the fields of speaking, reading, writing, and comprehension using a 5 point scale, which is an identical measurement method to that adopted in self-assessment. The subjects' English proficiencies are assessed in the same way by their present English native instructors. It is only the

subjects who score more than 3 points in at least 3 of the 4 skills in both languages that are to be claimed as 'balanced' bilinguals.

With respect to the reliability of teachers judgement in terms of models for assessment, problems are minimized as the Japanese native instructors teach both Japanese returnees and local Japanese students while the English native instructors have experienced teaching English to native speakers of English at a high school level.

2.2.4. SELECTED BALANCED BILINGUAL SUBJECTS

The results of balanced bilinguality measurements by C-/cloze-test, self-rating, and teacher judgements are shown in Table 2. (The self-rating data were collected through biographical data sheets (Appendix 3) which every participant filled in. This was done on a different occasion from the actual experiments, without any mention of the forthcoming experiments being made. The following data were gathered - onset age of L2 acquisition, LOR, strong language, language acquisition motivation, and self-ratings on Japanese and English. See Appendix 10 for a summary.)

Table 2: *Balanced bilingual subjects selection (total participants N=79)*

	C-test	cloze-test	self-rating (J)	self-rating (E)	teacher judgements (J)	teacher judgements (E)
passed	78	77	72	79	69	69
eliminated	1	2	7	0	10	10

* Some subjects failed across several measurement tests.

Of a total of 79 participants in the experiments, 65 scored satisfactorily in all the measurements to be qualified as 'balanced' bilinguals while 14 were unsatisfactory and eliminated. Of the 65 participants who passed, one was omitted due to her parents' dissent to include her data in this study, which makes the total subject number in the present research 64.

2.3. MATERIALS

The three word-lists used in Arnedt and Gentile's study (1986) were used in this study (Appendix 4). The first list of 16 words were used for picture-stimuli (taken from Peabody language development kits, level 1) and labelled in English, the second list of 16 Japanese words (translated from Arnedt and Gentile's original French) was translated into English, and the third list of 16 English words was copied. Some words were swapped among the original lists to avoid *katakana* (loan words from foreign languages) and *hiragana* as much as possible to examine the effects of the logographic features of Japanese (12 *kanji* and 4 *hiragana* appeared on the final list for the translation task). In place of these words with similar ratings for concreteness, ability to elicit an image, and familiarity (Toglia and Batting, 1978) were chosen to keep the average ratings constant among the three lists (see Appendix 4 for individual and mean ratings in each category & 3 examples of the slides). These were then randomly selected to make a 48-item list.

2.4. PROCEDURES

The experiments were conducted 6 times with 6 different grades during English classes by the instructors and in all cases in English. For ease of administration and to control the experimental conditions, the materials were presented on video using a Panasonic 29 inch TV. All the video-scripts are included in Appendix 5, but a brief outline of the video-taped instructions and item-presentation is given here. In the beginning, the subjects were informed of the general field this study is investigating and assured of confidentiality of the collected data. Then examiners were asked to distribute handout number 1 (Appendix 6) to the subjects (a one-minute pause was prerecorded onto the video with a still photo shown on the screen, so that the examiners need not pause or re-start the video machine when distributing handouts). Then the instructions of this experiment were given. The instructions used in Arnedt and Gentile's experiments (1986) were basically replicated here with a minor modification better suited to high school students. The modified version is:

> You will be shown fifty-three slides on a TV screen. Some will be pictures of very common objects, others of common Japanese words, and others will be of printed English words. When you see a picture, write down the English name for it on your sheet. When you see a Japanese word, translate it into English and write the English word on your sheet. When you see an English word on the screen, simply copy it onto your sheet. You will see a new slide every five seconds. The first five slides will give you a chance to become familiar with the amount of time between each slide and the speed at which you must write. So picture - write it in English, Japanese - translate into English, and English - just copy it on your sheet. Here are the trial five slides.

The five words used for the sample slides are also taken from Arnedt and Gentile's experiment (1986). They were randomly made up of two pictures (corresponding pictures are taken from Peabody language development kits), two English words, and 1 Japanese word (*kana*). Each slide was video-taped consecutively for five seconds controlled by the inbuilt digital timer of the Fujix-Hi8 FH120 video camera. A 5-second exposure duration per item has been endorsed to be sufficient enough to allow subjects to complete the encoding responses (Arnedt and Gentile, 1986; Matsumi, 1994). Having conducted the trial, subjects were shown the answers and allowed thirty seconds to ask any questions. Then the actual experiment began, showing on the TV screen 48 items randomly composed of pictures, English words, or Japanese words with a 5-second exposure time for each one. After four minutes, the instructors were asked to collect the handouts during a 30-second pause. Then handout number 2 (Appendix 7 A/B) was distributed for a Japanese cloze test and an English C-test, which took an additional 6 minutes following on from the instructions and sample test on the other side (side B) of the handouts which had already been completed. Upon completion of this task, the subjects were unexpectedly asked to recall as many as possible of the items that they wrote down during the encoding task. The exact instructions given were:

> This is the last but the most important section of this experiment. On the bottom half of side B, would you now please write down as many of the English words as possible that you wrote down when you were looking at the 48 slides regardless of whether they were pictures, Japanese, or English words ? For example, you saw 5 practice slides - 2 pictures of a comb and a chair respectively, one Japanese word "SAME', and 2 English words "man" and "hen". If you remember all, write down "comb" "chair" "shark" "man", and "hen" all in English. You have 5 minutes to recall and write them

down. You can start writing now.

A 5-minute pause took place on the TV screen during the incidental recall task. On finishing the task, the subjects were advised to return the handouts to the instructor and asked not to talk about the experimental procedures to their friends. The same experiments were to be conducted with different grades over a week long period. Finally, the instructors were asked to hand out sheet number 3 (Appendix 8) which were consent/dissent letters to the subjects' parents, stating the purpose of the experiment and the procedure to follow if they wanted their child's data to be excluded from the study [1]. The subjects were thanked for their cooperation. Prior to the experiments, all the instructors were explained the purpose of the experiment and shown the video as well as the scripts. Experimental conditions were kept constant across the 6 experiments to a fairly high degree. The material was presented entirely through the video.

1. This was a requirement of the Human Ethics Committee at Macquarie University.

3. RESULTS

Recall was scored individually against an intended words list. In doing so, each subject's written English encoding responses from the first phase of the experiment were also referred to as his or her input list. This was done because the labels assigned to pictures did not always correspond to their intended labels. Some slight variation also occurred in the English translations of Japanese words, since both 'ship' and 'boat', for instance, can serve as an English translation for the Japanese word 'FUNE'. Therefore responses on the recall list were defined as correct as long as they came under one of the following conditions:

(i) they matched the intended words:
e.g. whale - 'whale'

(ii) they matched the words on each subject's own input list:
- one of the translation forms possible for the Japanese word:
e.g. 'FUNE' - both 'ship' and 'boat' were marked correct
- totally different from the intended normative label, but the presented pictures could be interpreted as something else
e.g. fox - 'wolf' was marked as correct as long as 'wolf' was assigned to the same picture in his/her input list
- singular or plural forms of the normatives:
e.g. paws - paw; fox - foxes: both 'paw' and 'foxes' were marked as correct
- spelling mistakes:
e.g. pear - misspelled words such as 'pair' or 'pare' were marked as correct as long as same misspelled words were assigned to the same picture in his/her input list.

* The underlined words are the normatives. See Appendix 9 for more details.

The resulting overall means and standard deviations are presented in Table 3.1. and Figure 3.1 (a summary of the experimental results is given in Appendix 10 along with data on the subjects' biographical language information). Table 3.1. also displays mean recall scores according to various other categorizations of the sample, including gender, onset age of L2 acquisition, and length of residence in L2 speaking communities.

Table3.1 : Mean Words Recalled for Treatment Conditions by Subgroups

		Treatment Condition								
		Labelled Pictures			Translated Japanese			Copied English		
Subject Group	N	M	SD	RATIO	M	SD	RATIO	M	SD	RATIO
Total Sample	64	7.64	2.39	3.7	6.61	2.28	3.2	2.06	1.59	1.0
Gender										
Male	17	6.82	2.48	2.6	6.71	2.39	2.6	2.59	1.87	1.0
Female	47	7.94	2.31	4.2	6.57	2.26	3.5	1.87	1.45	1.0
Onset Age										
0-6 years old	33	8.00	2.17	4.9	6.18	2.14	3.8	1.64	1.48	1.0
over 7	31	7.26	2.58	2.9	7.06	2.37	2.8	2.52	1.61	1.0
Length of Residence										
0-4 years	18	7.67	2.45	3.0	6.33	1.91	2.5	2.56	1.42	1.0
5-10	38	7.32	2.39	3.9	6.89	2.45	3.7	1.87	1.70	1.0
over 11	8	9.13	1.81	4.9	5.88	2.23	3.1	1.88	1.36	1.0

Figure 3.1

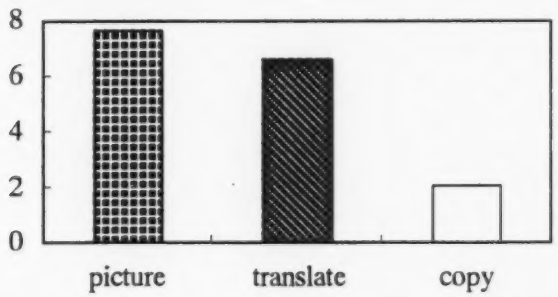


Figure 3.1: Incidental free recall scores (mean number of recalled English words that bilingual subjects generated by naming pictures, translating Japanese words, and copying English words)

The data show that the recall rate is highest for picture labels (mean recalled words = 7.64), next for the translation of Japanese words (6.61), and lowest for copied English words (2.06), although the first two experimental conditions show similar rates. Picture labels are recalled approximately 3.7 times as frequently as the copied English words, and translated Japanese words are recalled approximately 3.2 times as frequently as the copied English words. Thus the overall recall ratio for picture-labelling, translation, and copying conditions is 3.7 : 3.2 : 1.

The data in Table 3.1. were analyzed by two fixed factor analyses of variance with repeated measures on the last factor. In both analyses the last factor, encoding treatment (picture labelling *vs*. translation *vs*. copying), provides a within-subject comparison. The other factors provide between-subject comparisons. This analysis (a 2 X 3 ANOVA) was conducted first on the gender and encoding treatment, then on the LOR and encoding treatment, both of which revealed a nonsignificant interaction effect. However, the main effect for encoding treatment yielded a large and significant effect, $F(2, 59) = 111.75$, $p < .001$. A analysis of 2 X 3 ANOVA on onset age and encoding treatment showed a significant interaction effect, $F(2, 59) = 3.33$, $p < .05$. The final analysis of the data was completed using a 2 X 2 X 3 ANOVA with gender, onset age, and encoding treatment as the three independent variables. This

analysis yielded a nonsignificant interaction effect. Table 3.2. shows a summary of the interaction effects.

Table 3.2: Interaction effects on 3 variables and encoding treatment

variables	Encoding treatment conditions
Gender	nonsignificant
Onset age	**SIGNIFICANT**
Length of residence	nonsignificant

Thus significant differences were seen only (1) among encoding treatment conditions and (2) between onset age and encoding treatment. In order to determine exactly where the significance might be located, 3 paired t-Tests were conducted on 3 encoding conditions and 2 onset age groups (0-6, and 7+). As for the onset age group of 0-6 years, 3 t-Tests all showed a significant effect: (i) the picture-labelling and translation tasks, $t = 3.87$ ($df = 32$), $p < .05$, (ii) the picture-labelling and copying tasks, $t = 13.36$ ($df = 32$), $p < .05$, and (iii) the translation and copying task, $t = 8.73$ ($df = 32$), $p < .05$. On the other hand, a t-Test on the picture-labelling and translation task with the other group showed a nonsignificant effect while t-Tests on both the picture-labelling and copying tasks, and the translation and copying tasks showed significant effects with $t = 10.38$ ($df = 30$), $p < .05$ and $t = 10.23$ ($df = 30$), $p < .05$ respectively. These results are summarized in Table 3.3.

Table 3.3: Paired t-Tests on 2 onset age groups and encoding treatment

Onset age group	picture/translation	picture/copy	translation/copy

0 - 6 years old	significant	significant	significant
over 7	**NONSIGNIFICANT**	significant	significant

4. DISCUSSION

Based on the data, the four variables used in this experiment are examined for their possible involvement in the dual coding hypothesis, and discussed in the following order; (1) gender, (2) LOR, (3) onset age of L2 acquisition, and (4) linguistic specificity. Gender showed a nonsignificant effect either alone or in interaction with the other 3 encoding conditions, therefore it is only briefly discussed. LOR, which also showed a nonsignificant effect, is analyzed in relation to onset age. Finally the two variables of onset age of L2 acquisition and linguistic specificity are discussed.

4.1. GENDER

Gender showed no effect either alone or in interaction with the encoding conditions. This result is consistent with Paivio and Lambert's (1981), and seems to contribute to clarifying the gender issue in the dual coding hypothesis that there is no significant gender effect. However, broadening the scope a little wider than the bilingual dual coding theory might be useful before reaching a definite conclusion in this matter. Therefore general gender issues in other fields of bilingual studies are briefly reviewed.

Investigating a total of 305 Mexican-American bilingual

kindergarteners, Medina & Escamilla (1994) have reported no significant gender effects. A similar result was received by Kamada, King-Ying, Silva, and Noguchi (1995), who claimed that females have greater connectivity between the two halves of the brain providing more efficient second language acquisition. Contradictory findings, however, have been reported. Zentella (1990), for instance, looked at a very similar battery of subjects to the present study - 43 adolescent children of returnees to Puerto Rico, most of whom were raised in New York and who are English dominant speakers. By comparing this group with Puerto Ricans in New York, the researcher found that gender was the most relevant variable.

Researchers such as Saravan (1990), Cumming (1991), and Obler, Zatorre, Galloway, and Vaid (1982) have interpreted inconsistent findings in bilingual studies in terms of gender as unavoidable. They argue that gender *per se* may not play a major role, but when interaction occurs with handedness, age, onset age of L2 acquisition, L2 needs of learners, socioeconomic status or sociopsychological factors, gender may be affected. It seems that this argument explains the discrepant results reasonably well. Among all the available studies on the bilingual dual coding, Paivio and Lambert's Experiment 1 (1981) seems to be the only one that states the number of male and female subjects and explicitly discusses the issue of gender. In order to illuminate the gender issue clearly, the controlled variables of Paivio and Lambert's experiment and the present study are compared.

An early- *vs.* late-bilingual distinction was made in Paivio et al's experiment, where they defined early bilinguals as having acquired both languages in infancy and late bilinguals as acquiring 2 languages in childhood or early adolescence. This suggests that their L2 acquisition took place before early adolescence, which coincides with the subjects used in the present study in that they acquired L2 at ages 0 - 16 (but only one subject at age 15 and 16 respectively, and the rest before 13). Thus the subjects in both experiments match in onset age of L2 acquisition. Similarities are also seen in the nature of the bilingual subjects - balanced bilinguals. Applying multiple measurements to assess balanced bilinguality (a background questionnaire and a bilingual form of the Stroop speed of color-naming test by Paivio et al with C-/cloze-tests, self-ratings, and teacher judgements in this study), both experiments are presumed to have obtained high reliability in selecting balanced bilinguals.

As for age, Paivio et al's subjects were university students whereas the mean age of the subjects in the present study was 15.8, when the cognitive development is still on-going, although nearing its end according to Romaine (1995, p268). This means that age and maturity were unmatched in the two studies.

Thus it could be safely concluded that gender does not seem to play a major role in the bilingual dual coding theory as long as the theory is applied to balanced bilinguals who have acquired their L2 before early adolescence.

4.2. LOR

LOR showed a nonsignificant effect, $F(4, 120) = 2.34$, $p < .06$. However, this matter needs a brief exploration on the grounds that (1) the statistical analysis showed close proximity to the significant level of .05, and (2) some bilingual studies have found LOR effects on L2 acquisition (a summary is already given in 1.4.).

One way of analyzing the data would be to compare the results with what was predicted before the experiment. Based on Arnedt and Gentile's experimental results on manner of L2 acquisition (1986), this study hypothesized that the longer LOR is, the stronger the imagery-L2 connection ties would be, and the greater the imagery-translation difference would be in memory. This prediction would be verified by the results showing that the difference in number of mean recalled words between picture-labelling and translation conditions would be greater for the subjects who had stayed in L2 natural settings for longer periods of time. Table 4.2. summarizes the data (except for gender which was discussed above) collected either directly from the experiment or from the biographical language questionnaires.

Table 4.2: LOR

LOR years	N	Average LOR	Mean Age	Onset Age	*Back in Japan	Encoding Ratio	**Diff between P&T	***Motivation (E)	***Motivation (J)	****Self-ratings (E)	****Self-ratings (J)	Teacher Judgement (E)	Teacher Judgement (J)	Japanese Cloze-test	English C-Test
0-4	18	2.9	16.9	9.0	2.3	3.0:2.5:1.0	1.34	2.9	4.1	3.7	4.8	3.3	4	93%	88%
5-10	38	6.5	15.7	5.6	3.0	3.9:3.7:1.0	0.44	4.4	4.1	3.5	4.6	3.5	3.8	93%	99%
over 11	8	12.5	15.9	2.0	3.6	4.9:3.1:1.0	3.25	4.9	4.0	3.9	4.3	3.9	3.9	92%	100%

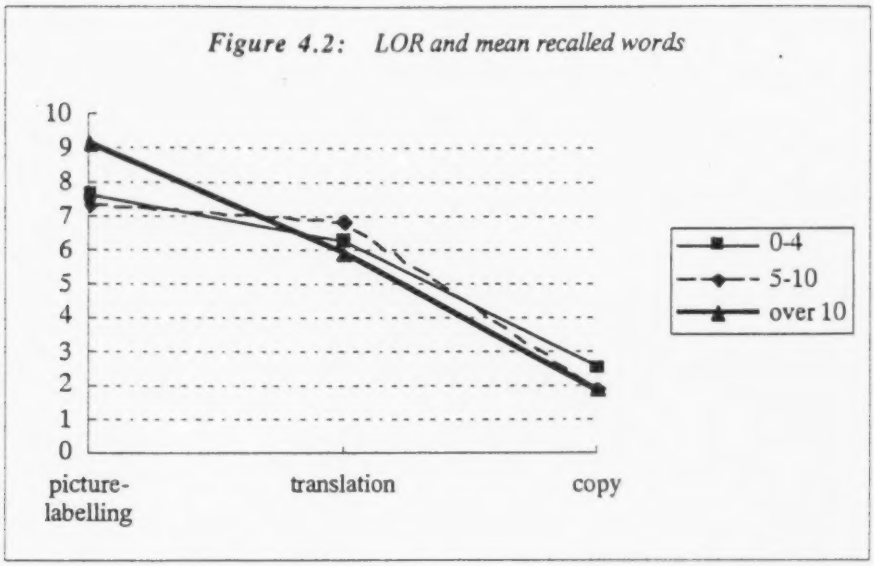
* Length of period after returning to Japan

** Difference in number of mean recalled words between picture-labelling and translation tasks in the experiment

*** Motivation is self-rated by each subject with '5' being very motivated. (E) means motivation of learning English when they were in English speaking communities. (J) means motivation of learning Japanese when they got back to Japan.

**** Self-ratings and teacher judgements are averaged across their ratings of reading, writing, speaking, and comprehension with '5' more superior to monolinguals and 1 more inferior.

Figure 4.2: LOR and mean recalled words



As shown in the column 'Diff between P & T' in Table 4.2. and Figure 4.2., the difference in the longest LOR group is greater than that in the shortest group as predicted. However, the LOR group of 5-10 years, whose difference was expected to lie somewhere between 3.25 and 1.34, results far below the shortest LOR group, that is, 0.44. To seek an explanation for this unexpected result, the other variables (in Table 4.2.) were examined to see whether similar discrepancies were observed with the LOR group of 5-10 years.

Mean ages and the length of time back in Japan exhibit virtually no difference across the three groups with mean ages from 15.7 to 16.0 and being back Japan from 2.3 to 3.16 years. Onset age of L2 acquisition

and LOR negatively correlate (r = -0.546 by Pearson correlational analysis), which indicates the consistent tendency that the younger they arrived in English speaking communities, the longer they stayed there. The motivation column seems to consistently suggest that the subjects are willing to study Japanese hard, irrespective of the difference in LOR, while the motivation to learn English was enhanced in the direction of the longer LOR rather than the shorter LOR. The self-ratings and teacher judgements columns appear to imply the consistent tendency that the subjects are apt to underestimate their Japanese proficiencies as they have stayed outside Japan for longer periods of time, while objective ratings by Japanese instructors are supported by the Japanese cloze-test results. As far as English proficiency is concerned, self-ratings and teachers judgements correlate positively with the shortest LOR group having the worst English proficiency and the longest the best, which is evidenced by the English C-Test results.

To sum up, the following could be induced:

(1) Present age, length of period back in Japan, motivation to learn Japanese, and Japanese proficiency do not seem to have caused the unpredicted results as they are kept the same across the three groups.

(2) Motivation to learn English, English self-ratings and English proficiency show a constant progression from the shortest LOR to the longest, while Japanese self-ratings show a constant regression from the shortest LOR to the longest. The LOR 5-10 group does not deviate from these two consistent tendencies, which seems to imply no relevance of

these variables to the discrepant or unexpected result in the LOR variable.

(3) Although LOR and onset age negatively correlate and the LOR 5-10 does not deviate from this tendency, the mean onset age of 5.6 for this group seems to need special attention due to the fact that age 5.6 approximately coincides with the age of beginning formal schooling in Japan, that is 6.

The age factor is to be discussed in the next section, however the cut-off ages in the next section are different in that the subjects are distinguished as either 0 - 6 or over 7 years old whereas the three LOR groups divide themselves into mean onset age groups of 2.0, 5.6, and 9.0 years old. Therefore the issue of onset age in this section is discussed in relation to schooling, particularly at the beginning.

In referring to education of minority children to avoid lagging behind their monolingual counterparts in academic achievement in both L1 and L2, Hamers and Blanc (1989) argue that 'there's strong evidence that promoting L1 literacy skills enhances overall academic achievement' (p206). In examining the relationship between age of arrival, length of residence and interdependence of literacy skills in English and Japanese among Japanese immigrant students in Toronto, Harley, Allen, Cummins,'and Swain (1990) found that continued development of academic skills in Japanese to a high level (i.e. comparable to that of students in the home country) was a difficult task for students who arrived in the host county at an early age, particularly

prior to formal schooling. Cummins (1991), who examined L1 maintenance and L2 acquisition of Chinese and Japanese students in North America, claim that L1 cognition and academic proficiency have a strong positive influence on academic success in English. Many other researchers such as Perozzi & Chavez-Sancher (1992) and Ricciardelli (1992) have found evidence supportive of the threshold theory proposed by Skutnabb-Kangas and Cummins (1988). The theory maintains that there may be levels of linguistic proficiency that bilingual children must attain to avoid cognitive deficits and to allow the cognitive benefits. It seems that this theory can possibly explain the results produced by the LOR group of 5-10 in this study, when integrated with the following two factors. The first is their mean age 5.6 (varying from 0 to 11) which indicates having received no formal schooling in L1. The second is the fact that approximately the moment the LOR 5-10 group arrived in the host countries, they started schooling entirely in L2 without any prior exposure to L2, where little L1 maintenance education was provided to support their cognitive development. The second factor is mostly verified by the data obtained through the language background questionnaire, which showed (1) that the subjects all attended English-speaking schools, (2) that only one of the 64 subjects was exposed to English prior to leaving Japan, and (3) only 30 % of the subjects attended Saturday Japanese schools.

One way of applying this argument to the bilingual dual coding theory would be that (1) L1-L2 translation ties within the LOR group of 5-10 did not develop as strongly as those by the LOR 0-4 group who had already established L1 to some extent through schooling prior to

their arrival in the host country at mean age 9.0 (varying from 0 to 16), and (2) starting school education in L2 without any L2 knowledge did not allow them to form L2-imagery system ties as strongly as those by the longest LOR group, because the longest LOR group had arrived in the host countries at mean age 3.25 (varying from 0 to 6), acquired English to a certain extent before school, and had already enhanced their L2-imagery ties. Both (1) and (2) are possibly due to an interaction between the subjects' underdeveloped L1 proficiencies with which the LOR group of 5-10 arrived in the host countries, and little promotion of L1 literacy skills in the L2 speaking communities. The combination of (1) and (2) might have caused a smaller difference between the picture-labelling and translation conditions than firstly predicted.

Another way of accounting for the unexpected results concerning the LOR group of 5-10 would be that under the circumstances in which subjects lived in the host countries, both L2-imagery and L1-L2 ties might have been equally strengthened. However the extent of this strengthening in comparison to L1-L2 ties of the shorter LOR groups and imagery-L2 ties of the longer LOR groups is indeterminate due to the lack of relevant data available. Despite this indeterminacy as to whether L1-L2 and imagery-L2 ties of the LOR group of 5-10 are equally loose or tight, the equal strength of connecting ties might have brought about the result of a smaller image-translation difference than previously predicted.

There is no way of knowing whether one of the above two speculations is correct or not without much more detailed personal data. However, the present discussion seems to indicate that even if LOR *per se* does not play a major role in the bilingual dual coding hypothesis, it could take an interactional role with the onset age of L2 acquisition, particularly in relation to the starting-school age. This interactional view could be supported in two ways. Firstly, further studies could test subjects with a different onset age to the one in this study and produce data showing that the imagery-translation difference in memory by the LOR 5-10 group lies within longer and shorter LOR groups. Secondly, subjects with LOR 5-10 and onset age of about 6 could be tested and produce similar results to this study.

To conclude, the imagery-translation difference was small for the shortest LOR group and it was great for the longest LOR as predicted. An unexpected result was produced in the LOR group of 5-10 years. It is inferred that this result is due to an interactional effect of the mean onset age of this group, which coincides with the beginning of formal schooling. More data, however, are needed to clarify this issue.

Despite this argumentation, however, the statistically nonsignificant effect of LOR remains unchanged, indicating that LOR does not play a major role in the bilingual dual coding theory.

4.3. ONSET AGE OF L2 ACQUISITION

Statistical analysis showed significant effects on tasks and onset age, $F(2,59) = 3.33$, $p < .05$. However t-Tests showed a nonsignificant effect on the picture-labelling and translation tasks with the onset age group of over 7 years old, while the remainder were all significant as seen in Table 3.3.

Table 3.3: Paired t-Tests on 2 onset age groups and encoding treatment

Onset age group	picture/translation	picture/copy	translation/copy
0 - 6 years old	significant	significant	significant
over 7	**NONSIGNIFICANT**	significant	significant

These two results are completely consistent with the predictions. The onset age difference is discussed first, followed by the issue of nonsignificant effects between picture-labelling and translation tasks with the older onset age group.

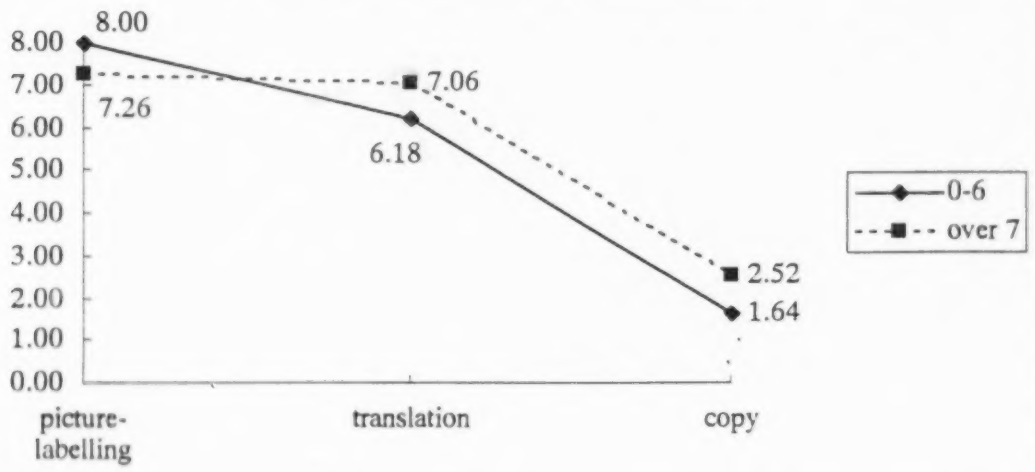
This study predicted that the earlier the onset age of L2 acquisition is, the stronger the L2-imagery ties become while the later the onset age is, the stronger the L1-L2 translation ties become. Based on this, the early onset age group was expected to show a greater difference in the number of recalled words between picture-labelling and translation conditions than the older group. As seen in Table 4.3.1. and Figure 4.3., the difference in recalled words between picture-

labelling and translation with the 0-6 onset age group is 1.92 words, which is greater than that of over 7 group, that is, .2. Thus the prediction on onset age seems to be supported.

Table 4.3.1: Onset age and number of recalled words in each condition

onset age	picture-labelling	translation	copy
0-6 years old	8.00	6.18	1.64
over 7	7.26	7.06	2.52

Figure 4.3: Mean numbers of recalled words



From a perspective of L2 acquisition manner, since both onset age groups have acquired their L2 in natural settings, they are presumed to produce a same degree of strong L2-imagery ties according to Arnedt

and Gentile (1986). This study, however, saw a difference in the two onset groups, which is accounted for from the standpoint of the type of formal schooling that took place.

The mean onset age for the onset group of 0-6 is 3.12 years old and that for the other group is 9.3. The mean LOR for the first group is 8.1 years and that for the second is 4.4. These data reveal that the younger onset group arrived in English-speaking countries at about age 3 and stayed there until they finished elementary school, while the other group arrived at 9 and stayed on until they were in grade 9. In other words, the younger group had about 3 years to acquire L2 (English) before starting formal schooling in L2, which subsequently lasted for about 6 years. Throughout this process, this group seems to have built up strong L2-imagery ties as the immersion group did in Arnedt and Gentile's experiments. On the other hand, the onset group of over 7 started their formal schooling in L1, but L1 education was only for 2 to 3 three years, then they were placed in L2 schooling without any prior exposure to L2. In receiving education in L2, this group seems to have developed strong L1-L2 translation ties based on their few years' education in L1 similar to Arnedt and Gentile's group who was taught L2 in core programmes. This reasoning seems to satisfactorily explain why there was a difference between the two onset age groups in the strength of L2-imagery ties, admitting that the argument was developed on a speculative basis using the mean age of each group.

Thus although the two onset age groups have both acquired L2 in natural settings, the language in which formal schooling started and the proficiencies of both L1 & L2 prior to schooling seem to interact in determining the strength of L1-L2 translation and L2-imagery connections. In this respect, Arnedt and Gentile's definition of 'manner' is subject to some modification or extension. Their definition of 'manner' only involved the way in which L2 is acquired, that is, whether L2 is learnt in natural settings (immersion programmes according to Arnedt and Gentile) or in formal education (core programmes). However, it should also include the 'manner' of how proficient L1 and L2 are prior to formal schooling, and also the 'manner' in which formal education takes place, and the sequence of languages, that is, L1 to L2 or L2 to L1. This extension seems to successfully explain both Arnedt & Gentile's and this study's results. Above all, this extension appears to be more generalizable across a wider range of age groups, particularly the group who are just before or after the initial stage of a formal education.

The focus of the discussion is now shifted onto the second issue of the nonsignificant effects between picture-labelling and translation conditions with the over-7-year-old group. The results of three major studies on the bilingual dual coding hypothesis and this study are summarized in Table 4.3.2. (the numbers indicate ratio, not mean recalled words, due to the unavailability of the mean number of recalled words in some studies).

Table 4.3.2: Comparison of ratio and difference of ratio in three experiments

experiment	subjects	*ratio	**difference in ratio
Paivio and Lambert (1981)	early/late	2.7 : 1.8 : 1	0.9
Arnedt & Gentile (1986)	core	3.1 : 2.4 : 1	0.7
	immersion	3.9 : 2.4 : 1	1.5
***Matsumi (1994)	****early bilingual	3.2 : 2.1 : 1	1.2
	*****late bilingual	2.9 : 2.0 : 1	0.9
	******L2 learner (beginners)	3.1 : 2.7 : 1	0.4
This study	onset age 0-6	4.9 : 3.8 : 1	1.1
	onset over 7	2.9 : 2.8 : 1	0.1

* ratio of mean number of recalled words under picture-labelling, translation, and copy conditions
** difference of ratio between picture-labelling and translation conditions
*** Matsumi's experiments were conducted under different conditions, therefore the figures are re-calculated for comparison purposes
**** early bilingual: acquired L2 as an infant
***** late bilingual: L2 learning started at age 12

Based on Arnedt and Gentile's findings (1986), this study predicted that the difference in the number of recalled words between picture-labelling and translation conditions with the early onset age group would be greater than that of the older group. The results show (Table 4.3.2.) that the ratio difference is 1.1 for the early group and 0.1 for the late, which is consistent with the prediction. Arnedt and Gentile's reasoning held true for their two subject types - L2 acquisition in core and immersion programmes. It also held true with Matsumi's three groups. His first group acquired L2 as infants, which strengthened L2-imagery ties according to Arnedt and Gentile, while late bilinguals and L2 learners (beginners) started learning L2 at age 12, resulting the L1-L2 ties being strengthened. The difference in ratio

of the early bilingual subjects between picture-labelling and translation modes was 1.1 while that of the late bilinguals was 0.9 and that of the L2 learners was 0.4. Thus the small difference between the picture-labelling and translation modes is consistent with Arnedt and Gentile's reasoning and this study's prediction.

With regards the degree of difference between the two encoding conditions, Matsumi's L2 learners and the subjects of onset age over 7 in this study statistically showed nonsignificant differences, whereas the core group in Arnedt and Gentile's experiment and Matsumi's late bilinguals showed significant effects. The reasons for which the first two groups showed nonsignificance are explored now.

The subjects in each experimental group and their mean age in the four bilingual dual coding studies are summarized in Table 4.3.3.

Table 4.3.3: Comparison of mean age in three experiments

experiment	subjects	mean age
Paivio & Lambert	early & late bilinguals	university students (not mentioned)
Arnedt & Gentile	core/immersion	university students (19.9 years old) + school teachers (30.2 years old)
Matsumi	early bilinguals	high school students(14-16 years old)
	late bilinguals	school teachers (27-42 years old)
	L2 learners (beginners)	high school students(14-16 years old)
This study	onset age 0-6	high school students (16.0 years old)
	onset over 7	high school students (16.0 years old)

* The subjects underlined indicate those who show nonsignificance in the difference

between picture-labelling and translation conditions.

One of the common features shared by the L2 learners in Matsumi's study and the older onset age group in this study is their age - high school students. In terms of L2 acquisition manner, the late bilinguals in Matsumi's study are validly recognized as older counterparts of the L2 learners in the same experiment in that their L2 acquisition began at age 12. In the same vein, the immersion group in Arnedt and Gentile's experiment seems similar to the onset age group of over 7 in this study (age apart), in that their bilinguality has not been acquired as infants but in natural settings. These two older counterparts - the late bilinguals in Matsumi's study and the immersion group in Arnedt and Gentile's study - have produced significant differences in the encoding conditions of picture-labelling and translation, while the younger have not. This could be interpreted that the differences between the two groups under discussion and their older counterparts are due to the age difference, that is, the fact that cognitive development of the younger subject groups is still on-going, while the older counterparts have already completed it.

If this reasoning, however, is true, it must also explain why the cognitive development has no effect on the onset age group of 0-6 in this study (mean age 16) and the early bilinguals in Matsumi's study (mean age 14-16). One way of interpretation would be due to the difference between the two onset age groups in terms of proficiencies in L1 and L2 prior to schooling and L1 maintenance as discussed in the previous section. Daller (1995), for example, examined the German and

Turkish language proficiency of Turkish students who emigrated to Germany at an early age and then returned to Turkey before finishing school, who are very similar to the subjects in this study in that they are also 'returnees'. In comparing the returnee students with the control group of students who remained in Turkey, the researcher found that because the returnees' L1 cognitive academic language proficiency was not fully developed before or during the acquisition of the second language, there was some difficulty in both languages. Therefore cognitive development in relation to language proficiencies of L1 and L2 and L1 maintenance could be a possible cause for the nonsignificant effects in this study.

Matsumi interprets the nonsignificant effects with his L2 learners as caused by the fact that L1-L2 ties have just begun to develop and have not developed as fully as his late bilingual subjects because the L2 learners started learning L2 only recently in core programmes. This is a plausible explanation for his subjects. However, using this explanation with the present study, the older onset age group should have shown a significant effect on the grounds that the subjects are already proficient balanced bilinguals who are supposed to have fully developed L1-L2 translation ties.

In investigating the English displacement effects (the positive correlation between age of English acquisition and ability in a non-English language) on bilinguals in America, Palij (1990) found that

(1) English proficiencies are the same for native speakers of English and those who acquired English as L2 before age 6, (2) their proficiencies are better than those who acquired it between ages 6-12 or later, and (3) there is about a 13-point drop in the SAT verbal score for every year of delay in the acquisition of English. He concluded that his findings might not only be due to the cognitive effects alone, but also due to interactions between cognitive processes and social processes, and/or due to interactions between cognitive processes and motivational factors. By social processes, the researcher meant 'societal pressures reinforcing English usage and removing support for the use of the non-English language' (p67), and by motivational factors 'to escape a non-English cultural background and assimilate into an English one' (p68). Taking these findings into account, it is possible to presume that social and motivational factors played a role in the results of this study.

Thus it seems that further research should be conducted to gather detailed data at least on cognitive, social, and motivational factors from subjects who have completed their cognitive development with onset age of L2 acquisition under control, before any conclusion on this matter may be reached.

This section dealt with the effects of L2 acquisition onset age on the bilingual dual coding theory. The experimental results were consistent with the predicted difference in age and smaller differences between picture-labelling and translation encoding conditions. In the

process of explaining this consistency, it has been revealed that not only may the manner of L2 acquisition (Arnedt and Gentile, 1986) play an important role in the bilingual dual coding theory, but also proficiencies of L1 and L2 prior to schooling, and the language sequence in formal education (L1 to L2 or L2 to L1). As for the nonsignificant differences between picture-labelling and translation conditions with the older onset age group, cognitive development along with social and motivational factors were discussed as plausible causes, however more data are necessary to reach a firmer conclusion.

4.4. LANGUAGE SPECIFICITY

The overall 3.7 : 3.2 : 1.0 ratio for picture-labelling : translation : copy conditions in this experiment is consistent with the predictions (1) that an approximate ratio of 3 : 2 : 1 would be produced in the bilingual dual coding theoretical framework, and (2) the translation condition would produce a closer ratio to the imagery condition due to the logographic features of *kanji*. Individual predictions are to be closely examined in this section.

The overall ratio for pictorial : translation : copy encoding conditions in this study is compared with those of other bilingual dual coding studies in Table 4.4.

Table 4.4: Comparison of ratios, differences, and languages

Experiments	Languages in bilingual subjects	Ratio	*Differences
Paivio & Lambert (1981)	English/French	2.7 : 1.8 : 1.0	0.9
Arnedt & Gentile (1986)	English/French(core)	3.1 : 2.4 : 1.0	0.7
	English/French(immersion)	3.9 : 2.4 : 1.0	1.5
Matsumi (1994)	Japanese/German(early bilingual)	3.2 : 2.1 : 1.0	1.1
	Japanese/English(late bilingual)	2.9 : 2.0 : 1.0	0.9
This study	Japanese/English	3.7 : 3.2 : 1.0	0.5

* differences in ratio between pictorial and translation encoding conditions

While the ratio of the pictorial condition against the copy condition among the three previous studies ranges from 2.7 (Paivio & Lambert) to 3.9 (Arnedt & Gentile's core group), the present study shows 3.7. Arnedt and Gentile interpreted the high superiority for picture recall (ratio of 3.9 against the copy conditions) in their experiments as being 'in accordance with the findings of Paivio and Lambert' (p297) and 'supportive of dual coding theory for bilingual memory' (p298). In line with this argument, the ratio 3.7 in the present study seems to well fit in the bilingual dual coding framework.

Regarding the ratio of the translation condition against the copy condition, this study shows 3.2, which is higher than any other studies, although the statistical analysis showed significant differences among the three encoding treatment conditions. Before discussing whether this result supports the dual coding hypothesis or not, language specificity is

dealt with first, because it was predicted that the high translation ratio might be caused by Japanese language specificity.

Due to the language specificity - logographic features of Japanese *kanji* -, the pictorial/translation difference in Japanese/English bilingual memory was predicted to be smaller than that in Roman-alphabetic bilinguals. As Table 4.4. shows, the ratio differences in English/French bilinguals range from 1.5 (Arnedt & Gentile's immersion group) to 0.7 (their core group), while Japanese/English bilinguals in the present experiment show 0.5. This result is consistent with the prediction and seems to lend support to the hypothesis of language specificity.

Matsumi's Japanese/alphabetic bilingual subjects, however, show ratio differences of 0.9 to 1.1, which is not greatly different from alphabetic bilinguals. This difference in ratio between Matsumi's and the present study could be derived from their elicitation methods. This study adopted the same encoding conditions of picture-labelling, translation, and copying as Paivio & Lambert (1981) and Arnedt & Gentile (1986). On the other hand, Matsumi required the subjects to translate or copy a mixed list of L1 and L2 words either under imagery (visualize the presented word that came to mind) or non-imagery conditions. Thus Matsumi's subjects were presented with either L1 words to translate or L2 words to copy, not with pictures at all. The process of visualizing a stimulus to mind under the imagery condition could be interpreted as requiring more of an effort by the subjects to generate both an image and translation. This effect could have caused the generated words to remain in memory comparatively longer, which

resulted in a larger number of recalled words in translation under the imagery condition than translation under the non-imagery condition. This explains the great difference between translation under imagery and non-imagery conditions in Matsumi's experiment.

The difference between Matsumi's and this study is now examined. Matsumi's translation mode under the imagery condition seems equivalent to the picture-labelling mode in the present study in that L1 & L2 verbal and imagery systems are all activated. Compared with the straightforward translation of the imagery stimulus into L2 in this study, Matsumi's imagery translation elicitation seems to have needed to exert more effort in generating both a picture and translation. More effort, therefore longer retention in memory, in Matsumi's imagery translation apparently seems to help generate more recall words than the pictorial mode in the present study. Combining this reasoning with the fact that Matsumi's non-imagery translation mode and the translation mode in this study are identical in that both involve only L1 and L2 verbal systems, seems to account for the greater imagery-translation difference in Matsumi's experiment than the present study.

Despite the satisfactory explanation as to why Matsumi's results were not consistent with this study's results, it still appears more appropriate to compare this study directly with Paivio & Lambert's and Arnedt & Gentile's studies rather than Matsumi's experiment on the

grounds that elicitation methods should be kept the same for comparison across various experiments. Thus the predicted smaller difference in the ratio of pictorial-translation modes in logographic/alphabetic bilinguals seem to be verified when compared with Roman-alphabetic bilinguals.

The similar processing of logographic language (*kanji* and Chinese characters, in particular) and pictures, which results in a strong memory retention, are reported by a number of researchers such as Chen & Juola (1982), Tzeng & Wang (1983), Paivio (1986), Ho & Chen (1993), Leck, Weekes & Chen (1995), and Brosig (1993). Similarities are also supported from a number of neuropsycholinguistic studies which claim both *kanji* and pictures are processed in the right cerebral hemisphere while non-logographic Japanese, *kana*, and alphabetic languages are processed in the left (e.g. Hatta, 1977, 1981; Morikawa, 1981; Nguy, Allard, and Bryden, 1980).

With sufficient support for the logographic effects (similar nature to pictures), the smaller pictorial-translation differences in this study seems to be supported, which at the same time would validate the ratio 3.2 for the translation condition in this study. Thus the 3.7 : 3.2 : 1.0 ratio for the encoding treatment could be interpreted as supportive of the bilingual dual coding hypothesis in the following way.

The recall for pictures is more than 3 times that of the copying

condition, therefore supportive of the notion that the imagery mode is the most accessible to recall. Recalling more than three times (due to the language specificity, thus interpreted as 'more than twice' without the effects) that of the copied items, the translation mode supports the notion of two independent verbal systems which help recall in an additive fashion. Both picture-labelling and translation tasks are presumed to involve activation and storage of independent memory representations with additive effects on the likelihood of retrieving items. The resultant ratio of 3.7 : 3.2 seems to lend support to this presumption when Japanese language specificity is taken into account. In other words, the results from the pictorial task could be seen to reveal that L1 and L2 verbal systems and the common imagery system were activated, and were thus 3 times able to recall and retrieve items later as the copy task, where it is presumed that only the L2 verbal system is activated. Similarly, the translation task seems to have shown that the task involved both L1 and L2 verbal systems, and was thus more than twice as memorable as the unilingual task of copying. As Paivio and Desrochers suggested (1980), the additive effects observed in this experiment seems to imply that two verbal systems and a pictorial system function independently, while the involvement of the three systems in the picture-labelling task and the two verbal systems in the translation task seem to imply that the three systems are interconnected. Thus the results in this study are completely consistent with predictions from the bilingual dual coding hypothesis when the language specificity is considered, which enhanced the ratio for the translation mode in the present study.

5. CONCLUSION

The results of this experiment - a 3.7 : 3.2 : 1 ratio for imagery : translation : copy - seem supportive of the bilingual dual coding hypothesis proposed by Paivio and Desrochers (1980). The additive effects are supportive of the independence of two verbal systems and a common imagery system, whereas the activation of three systems in the picture-labelling task and 2 systems in the translation task lend support to an interconnectedness of the three systems. The closer ratio of the picture-labelling and translation modes as well as the high ratio of the translation task were interpreted as caused by specific Japanese language features - logographic effects. No gender effects were observed. Although LOR showed nonsignificant effects on the encoding treatment, the longest LOR group produced a larger imagery-translation difference than the shortest group as predicted. The group in between them was presumed to have impacted on the nonsignificant results in terms of their onset age of L2 acquisition, which coincided with the age when formal schooling starts in Japan - age 6. Through discussion on onset age, it was found that not only the manner of L2 acquisition (as proposed by Arnedt & Gentile, 1986), but also proficiencies of L1 and L2 prior to formal schooling and the language shift in school education (whether L1 to L2 or L2 to L1) might play a role in forming L1-L2 or L2-imagery ties. Onset age of L2 acquisition was found to be significant, that is, the earlier it is, the stronger the L2-imagery system connecting ties are enhanced.

In summary, the present study seems to provide evidence that the bilingual dual coding theory is generalizable across (1) bilinguals not only in Indo-European alphabetic languages, but also in alphabetic/non-alphabetic languages, and (2) bilingual in their childhood as well as adulthood.

This study has raised two problematic factors, which need to be resolved in future studies. Firstly, as Hamers and Blanc (1989, p22) maintain, 'the development of tests designed to capture the bilingual's specific competence is an urgent task for researchers'. Since the bilingual dual coding theory is based on balanced bilinguals, the first step for researchers exploring this theory is to select appropriate subjects. Without tests to measure bilingual balance, they can never be certain of their findings even if they adopt manifold assessment methods as done in this study. Multiple assessments are the best researchers can do at present, however development of proper tests is urgently needed to validate obtained data in terms of the appropriateness of subjects.

Secondly, in analyzing data obtained in this study, it turned out that there has been only scant research conducted on the bilingual dual coding hypothesis, despite its prospective theory on bilingual memory. This has led the researcher to bring in the findings from a wider scope of other bilingual studies in comparison with the results of this study. Furthermore, even the very limited number of studies on the bilingual

dual coding hypothesis differ in their variables and methodologies, which made direct comparison of the results and the findings of this study, difficult. Thus further research on this theory is necessary to accumulate more data to enable the theory to encompass a wider bilingual population or allow greater generalizability. It cannot be emphasized too much, in conducting further studies that the controlling variables are crucial since bilinguality is known to involve many variables such as socioeconomic status, cognitive development (skills), personality, sociolinguistic proficiency, and motivation (Romaine, 1995, p273). Cognitive aspects have received special attention in this study, which used high school students as the subjects. Even an analysis from the perspective of parents' attitudes towards education and life could have foregrounded new findings, as Romaine claims that their attitudes could play some role in the bilinguality of their children (1995, p279). Thus in studies where children are involved as subjects, a more cautious control of variables would seem desirable in future studies.

REFERENCES

Arnedt, C. S. and **Gentile, J. R.** 1986. A Test of Dual Coding Theory for Bilingual Memory. *Canadian Journal of Psychology*, **40** (3), 290-299.

Brosig, E. 1993. Bilingual Linguistic Memory in Reference to the Mode of Encoding. *Italiana di Linguistic Applicata*, **25** (1), 99-111.

Bugelski, B.R. 1970. Words and things and images. *American Psychologist*, **25**, 1002-1012.

Cabeza, R. 1995. Investigating the mixture and subdivision of perceptual and conceptual processing in Japanese memory test. *Memory & Cognition,* **23** (2), 155-165.

Caramazza, A., and **Brones, I.** 1980. Semantic classification by bilingual. *Canadian Journal of Psychology*, **34**, 77-81.

Chen, H. C. and **Juola, J. F.** 1982. Dimensions of lexical coding in Chinese and English. *Memory and Cognition*, **10**, 216-224.

Cohen, A.D. 1980. *Testing Language ability in the classroom*. Massachusetts: Newbury House Publishers.

Cummins, J. 1984. *Bilingualism and special Education: Issues in Assessment and Pedagogy.* Clevedon, Avon: Multilingual Matters.

Cummins, J. 1991. Language Learning and Bilingualism. *Sophia Linguistica*, **29**, 1 - 194.

Daller, H. 1995. The Academic Language Proficiency of Turkish Returnees from Germany. *Language culture and curriculum*, **8** (2), 163-173.

Dalrymeple-Alford, E. C., and **Aamiry, A.** 1969. Language and category clustering in bilingual free recall. *Journal of Verbal*

Learning and Verbal Behavior, **8**, 762-768.

De Groot, A. M. B. 1993. Word-Type Effects in Bilingual Processing Tasks: Support for a Mixed-Representational System. In Schreuder & Weltens (eds.). *The bilingual lexicon*. 27 -52. Philadelphia: John Benjamins Publishing Co.

Durgunoglu, A. Y., and **Roediger, H. L.** 1987. Test differences in Accessing Bilingual Memory. *Journal of memory and Language*, **26**, 377-391.

Groot, A.M.B. and **Hoeks, J.C.J.** 1995. The Development of Bilingual Memory: Evidence from Word Translation by Trilinguals. *Language Learning*, **45** (4), 683-724.

Hamers, J.F. and **Blanc, M. H. A.** 1989. *Bilinguality and Bilingualism.* Cambridge: Cambridge University Press.

Harbluk, J., Paivio, A., and **Clark, J.M.** 1987. Bilingual dual coding theory: Differences between translations and synonyms as a function of concreteness. Unpublished manuscript. In Paivio, A., Clark, J.M., and Lambert, W.E. 1988. Bilingual Dual-Coding Theory and Semantic Repetition Effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14** (1), 163-172.

Harley, B., Allen, P., Cummins, J., and **Swain, M.** (eds) 1990. *The Development of Second Language Proficiency.* Cambridge: Cambridge University Press.

Hatta, T. 1977. Lateral Recognition of Abstract and Concrete KANJI in Japanese. *Perceptual and Motor Skills*, **45**, 731-734.

Hatta, T. 1981. Differential Processing of Kanji and Kana stimuli in Japanese people: Some Implications from Stroop-Test Results. *Neuropsychologia*, **19**, 87-93.

Heaton, J.B. 1988. *Writing English Language Tests*. New York: Longman.

Hirose, T. 1992. Recognition of Japanese KANA Words in Priming Tasks.

Perceptual and Motor Skills, **75**, 907-913.

Ho, C.S. and **Chen H.** 1993. Effects of Syntactic Structure in the Memory of Concrete and Abstract Chinese Sentences. *Journal of Psycholinguistic Research,* **22** (5), 505-518.

Hoffmann, C. 1991. *An introduction to bilingualism*. New York: Longman.

Hood, M.A.G. 1990. The C-Test: A viable alternative to the use of the cloze procedure in testing?. In *Language Proficiency: Defining, Teaching, and Testing. Arena*, L.A. (ed.). New York: Plenum Press.

Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Hummel, K. M. 1993. Bilingual memory research: From storage to processing issues. *Applied Psycholinguistic*, **14**, 267-284.

Johnson, J. and Newport, E. 1991. Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition*, **39**, 215-258.

Kamada, L., Xing-Ying, L., Silva, W., and **Noguchi, M. G**. 1995. The Nurture and Nature of Bilingual Acquisition.In Troyer, Cornwell, and Morikawa (eds). On JALT 95: *Curriculum and Evaluation*. Tokyo: JALT.

Kirsner, K., Lalor, E., and **Hird, K.** 1993. The Bilingual Lexicon: Exercise, Meaning and Morphology. In Schreuder & Weltens (eds.). *The bilingual lexicon*. 215 -248. Philadelphia: John Benjamins Publishing Co.

Klein-Braley, C. 1985. A cloze-up on the C-Test: a study in the construct validation of authentic test. *Language Testing*, **2** (1), 76-104.

Kolers, P. 1963. Interlingual word association. *Journal of Verbal Leaning and Verbal Behavior*, **2**, 291-300.

Kolers, P. 1965. Bilingualism and bicodalism. *Language and Speech*, **8**, 122-128.

Kolers, P. 1966a. Interlingual facilitation of short-term memory.*Journal of Verbal Leaning and Verbal Behavior,* **5**, 314-319.

Kolers, P. A. 1968. Bilingualism and Information Processing. *Scientific American,*78-87.

Kolers, P.A. and **Gonzalez, E.** 1980. Memory for Words, Synonyms, and Translations. *Journal of Experimental Psychology: Human Learning and Memory*, **6** (1), 53-65.

Krashen, S. D. 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

Kroll, J. F. 1993. Accessing conceptual representations for words in a second language, in Schreuder, R., and Weltens, B. (eds.) *The Bilingual Lexicon*. Philadelphia: John Benjamins Publishing Co.

Lambert, W.E., Ignatow, M., and **Krauthamer, M.** 1968. Bilingual organization in free recall. *Journal of Verbal Learning and Verbal Behavior*, **7**, 207-214.

Lambert, W. E. and **Rawlings, C.** 1969. Bilingual Processing of Mixed-Language Associative Networks. *Journal of Verbal Learning and Verbal Behavior,* **8**, 604-609.

Leck, K. J., Weekes, B. S., and Chen, M. J. 1995, Visual and phonological pathways to the lexicon: Evidence from Chinese readers. *Memory and Cognition*, **23** (4), 468-476.

Macleod, C.M. 1976. Bilingual episodic memory: Acquisition and forgetting. *Journal of Verbal Learning and Verbal Behavior*, **15**, 347-364.

Macnamara, J. 1967. The bilingual's linguistic performance: A psychological overview. *Journal of Social Issues,* **23**, 58-77.

Madsen, H.S. 1983. *Techniques in Testing*. New York: Oxford University Press.

Matsumi, N. 1994 Processes of words memory in second language acquisition: A test of bilingual dual coding theory. *The Japanese Journal of Psychology,* **64** (6), 460-468.

Median, M. and **Escamilla, K.** 1994. Language Acquisition and Gender for Limited-Language-Proficient Mexican American in a Maintenance Bilingual Program. *Hispanic Journal of Behavioral Science*, **16** (4), 422-437.

Mgistre, E. 1992. Second Language Learning in Elementary and Hish School Studies. *European Journal of Cognitive Psychology*, **4** (4), 355- 365.

Morikawa, Y. 1981. Stroop Phenomena in the Japanese Language: The Case of Ideographic Characters (KANJI) and Syllabic Characters (KANA). *Perceptual and Motor Skills*, **53**, 67-77.

Morton, J. 1979. Facilitation in word recognition: experiments causing change in the logogen model. In Kokers, Wrolstad, and Bouma (eds.) *Processing of Visible Language*. New York: Plenum Press.

Newport, E. 1990. Maturational Constraints on Language Learning. *Cognitive Science*, **14**, 11-28.

Nguy, T. V. H., Allard, F. A., and Bryden, M. P. 1980. Laterality Effects for Chinese Characters: Differences between Pictorial and Nonpictorial Characters. *Canadian Journal of Psychology*, **34** (3), 270-273.

Nott, C.R., and **Lambert, W.E.** 1968. Free recall of bilinguals. *Journal of Verbal Learning and Verbal Behavior*, **7**, 1065-1071.

Obler, L., Zatorre, R., Galloway, L., and **Vaid, J.** 1982. Cerebral Lateralization in Bilinguals: Methodological Issues. *Brain and Language*, **15** (1), 40-54.

Oller, J. W. 1979. *Language Test at School.* London: Longman.

O'neill, W., Roy, L., and **Tremblay, R.** 1993. A translation-based generation effect in bilingual recall and recognition. *Memory and Cognition*, **21** (4), 488-495.

Ono, H. 1994. *Science of bilingual* (title in Japanese; Bairingaru no kagaku). Tokyo: Kodansha.

Oyama, S. 1976. A sensitive period in the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research*, **5**, 261-285.

Paivio, A. 1971. *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Paivio, A. 1965. Abstractness, imagery and meaningfulness in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, **4**, 32-38.

Paivio, A. 1975. Coding distinctions and repetition effects in memory. In G.H. Bower (Ed.), *The psychology of learning and motivation*, **9**. New York: Academic Press.

Paivio, A. 1983. The Empirical Case for Dual Coding. In J. Yuille (ed.), *Imagery, Cognition and Memory*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Paivio, A. 1986. *Mental representations: A dual-coding approach.* New York: Oxford University Press.

Paivio, A. 1991. *Images in mind: The evolution of a theory.* London: Harvester Wheatsheaf.

Paivio, A. and **Csapo. K.** 1973. Pictures superiority in free call: Imagery or dual coding? *Cognitive Psychology*, **5**, 176-206.

Paivio, A., Clark, J.M., and **Lambert, W.E.** 1988. Bilingual Dual-Coding Theory and Semantic Repetition Effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14** (1), 163-172.

Paivio, A., and **Desrochers, A.** 1980. A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, **34**, 388-399.

Paivio, A. and **Lambert, W.** 1981 Dual Coding and Bilingual Memory. *Journal of Verbal Learning and Verbal Behavior,* **20**, 532-539.

Palij, M. 1990. Acquiring English at Different Ages: The English Displacement Effect and Other Findings. *Journal of Psycholinguistic Research*, **19** (1), 57-69.

Paradis, M. (ed.) 1978. *Aspects of Bilingualism*. Columbia, SC: Hornbeam Press.

Paradis, M. 1981. Neurolinguistics organization of a bilingual's two languages. In Copeland, J. ed., *The Seventh LACUS Forum*. Columbia, SC: Hornbeam Press.

Paradis, M. and **Lebrun, Y.** 1984. Early bilingualism and child development. *Neurolinguistics*, **13**, 9-18.

Peabody language development kits, level 1. 1965. Minnesota: American Guidance Service, INC.

Perozzi, J. and **Chavez-Sanchez, M.** 1992. The effect of Instruction in L1 on Receptive Acquisition of L2. *Language, Speech, and Hearing Services in Schools*, 23 (4), 348-352.

Piper, A. 1983. A Comparison of the Cloze and C-test as placement test items. *British Journal of Language Teaching* vol.**21** (1), 45-50.

Potter, M. C., So, K., von Eckhardt, B., and **Feldman, L.B.** 1984. Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, **23**, 23-38.

Ricciardelli, L. 1992. Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research*, **21** (4), 301-316.

Romaine, S. 1995. *Bilingualism*. Cambridge: Blackwell Publishers Inc.

Saravan, V. 1990. Gender Difference in a Bilingual Tamil-English Community. *Australian Review of Applied Linguistics*, **13** (2), 97-113.

Scarborough, D.L., Gerard, L., and **Cortese, C.** 1984. Independence of Lexical Access in Bilingual Word Recognition. *Journal of Verbal Learning and Verbal Behavior*, **23**, 84-99.

Scovel, T. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Philadelphia: Newbury House Publishers.

Shohamy, E. 1983. The stability of the oral proficiency trait on the oral interview speaking test. *Language Learning,* **33**, 527-540.

Skkutnabb-Kangas, T. and **Cummins, J.** 1988. *Minority Education: from shame to struggle*. Clevedon: Multilingual Matters.

Singleton, D. and **Little, D.** 1991. The second language lexicon: some evidence from university-level learners of French and German. *Second Language Research,* **7** (1), 61-81.

Snodgrass, J.G. 1984. Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior*, **23**, 3-22.

Snodgrass, J. G. 1993. Translating versus Picture Naming: Similarities and Differences. In Schreuder & Weltens (eds.). *The bilingual lexicon*. 83 -114. Philadelphia: John Benjamins Publishing Co.

Snodgrass, J. G. 1995. Organization of the bilingual lexicon: Categorical versus alphabetic cuing in Russian-English bilinguals. *Journal of Psycholinguistic Research,* **24** (2),145-163.

Snodgrass, J.G., and **McClure, P.** 1975. Storage and retrieval properties of dual codes for pictures and words in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, **1**, 521-529.

Schwanenflugel, P.J., and **Rey, M.** 1986. Interlingual Semantic Facilitation:Evidence for a Common Representational System in the Bilingual Lexicon. *Journal of memory and languages*, **25**, 605-618.

Taylor, I. 1971. How are words from two languages organized in bilinguals' memory?. *Canadian Journal of Psychology,* **23**, 228-240.

Tyson, G.A., Doctor, E.A., and **Mentis, M.** 1988. A Psycholinguistic Perspective on Bilinguals' Discrepant Questionnaire Responses. *Journal of cross-cultural psychology*, **19** (4), 413-426.

Tzeng, O. J. L. and **Wang, W. S. Y.** 1983. The first tow R's. *American Scientist*, **71**, 238-243.

Vaid, J. 1988. Bilingual Memory Representation: A Further Test of Dual Coding Theory. *Canadian Journal of Psychology*, **41** (1), 84-90.

Yumoto, K. 1995. Age Factors and Language Proficiency in Child SLA. In Troyer, Cornwell, and Morikawa (eds). On JALT 95: *Curriculum and Evaluation*. Tokyo: JALT.

Watkins, M.J., and **Peynircioglu, Z.F.** 1983. On the nature of word recall: Evidence for linguistic speficity. *Journal of Verbal Learning*